

(19)



Europäisches Patentamt
European Patent Office
Office européen des brevets



(11)

EP 0 953 639 A1

(12)

EUROPEAN PATENT APPLICATION

(43) Date of publication:
03.11.1999 Bulletin 1999/44

(21) Application number: 98107925.4

(22) Date of filing: 30.04.1998

(51) Int. Cl.⁶: **C12N 15/13**, C07K 16/40,
C07K 16/46, C12N 15/62,
C12N 15/85, C12N 5/10,
C07K 19/00, A61K 47/48,
A61K 51/10, A61K 39/395,
G01N 33/577, G01N 33/574

(84) Designated Contracting States:
**AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU
MC NL PT SE**
Designated Extension States:
AL LT LV MK RO SI

(71) Applicant:
**Boehringer Ingelheim International GmbH
55216 Ingelheim (DE)**

(72) Inventors:
• **Park, John Edward
88400 Biberach/Riss (DE)**

- **Garin-Chesa, Pilar
88400 Biberach/Riss (DE)**
- **Bamberger, Uwe
88416 Ochsenhausen (DE)**
- **Leger, Olivier
74100 Annemasse (FR)**
- **Saldanha, Jose
Enfield Middlesex, EN1 1TE (GB)**
- **Rettig, Wolfgang J.
88400 Biberach a.d. Riss (DE)**

Remarks:

The applicant has subsequently filed a sequence
listing and declared, that it includes no new matter.

(54) **FAPalpha-specific antibody with improved producibility**

(57) Recombinant antibody proteins are provided
that specifically bind fibroblast activation protein alpha
(FAP α) and comprise framework modifications resulting
in the improved producibility in host cells. The invention
also relates to the use of said antibodies for diagnostic
and therapeutic purposes and methods of producing
said antibodies.

EP 0 953 639 A1

Application No. 10/643,857
Attorney Docket No. 13783-105022
References
B14

Description**Field of the invention**

5 **[0001]** The present invention relates to antibody proteins that specifically bind fibroblast activation protein alpha (FAP α). The invention also relates to the use of said antibodies for diagnostic and therapeutic purposes and methods of producing said antibodies.

Background of the invention

10 **[0002]** The invasive growth of epithelial cancers is associated with a number of characteristic cellular and molecular changes in the supporting stroma. A highly consistent molecular trait of the reactive stroma of many types of epithelial cancer is induction of the fibroblast activation protein alpha (from now on referred to as FAP), a cell surface molecule of reactive stromal fibroblasts originally identified with monoclonal antibody F19 (Garin-Chesa P., Old L. J. and Rettig W. J. (1990) Cell surface glycoprotein of reactive stromal fibroblasts as a potential antibody target in human epithelial cancers. *Proc. Natl. Acad. Sci.* **87**: 7235). Since the FAP antigen is selectively expressed in the stroma of a range of epithelial carcinomas, independent of location and histological type, a FAP-targeting concept has been developed for imaging, diagnosis and treatment of epithelial cancers and certain other conditions. For this purpose a monoclonal antibody termed F19 that specifically binds to FAP was developed and described in US Patent 5,059,523, which is hereby
20 incorporated by reference in its entirety.

[0003] One serious problem that arises when using non-human antibodies for applications *in vivo* in humans is that they quickly raise a human anti-non-human response which reduces the efficacy of the antibody in patients and impairs continued administration. Humanisation of non-human antibodies is commonly achieved in one of two ways: (1) by constructing non-human/human chimeric antibodies, wherein the non-human variable regions are joined to human constant regions (Boulianne G. L., Hozumi N. and Shulman, M. J. (1984) Production of functional chimaeric mouse/human antibody *Nature* **312**: 643) or (2) by grafting the complementarity determining regions (CDRs) from the non-human variable regions to human variable regions and then joining these "reshaped human" variable regions to human constant regions (Riechmann L., Clark M., Waldmann H. and Winter G. (1988) Reshaping human antibodies for therapy. *Nature* **332**: 323). Chimeric antibodies, although significantly better than mouse antibodies, can still elicit an anti-mouse response in humans (LoBuglio A. F., Wheeler R. H., Trang J., Haynes A., Rogers K., Harvey E. B., Sun L., Ghraieb J. and Khazaeli M. B. (1989) Mouse/human chimeric monoclonal antibody in man: Kinetics and immune response. *Proc. Natl. Acad. Sci.* **86**: 4220). CDR-grafted or reshaped human antibodies contain little or no protein sequences that can be identified as being derived from mouse antibodies. Although an antibody humanised by CDR-grafting may still be able to elicit some immune reactions, such as an anti-allotype or an anti-idiotypic response, as seen even with natural human antibodies, the CDR-grafted antibody will be significantly less immunogenic than a mouse antibody thus enabling a more prolonged treatment of patients.

[0004] Another serious limitation relating to the commercial use of antibodies for diagnosis, imaging and therapy is their producibility in large amounts. In many instances recombinant expression of native, chimeric and/or CDR-grafted antibodies in cell culture systems is poor. Factors contributing to poor producibility may include the choice of leader sequences and the choice of host cells for production as well as improper folding and reduced secretion. Improper folding can lead to poor assembly of heavy and light chains or a transport incompetent conformation that forbids secretion of one or both chains. It is generally accepted, that the L-chain confers the ability of secretion of the assembled protein. In some instances multiple or even single substitutions can result in the increased producibility of antibodies.

[0005] Because of the clinical importance of specific immunological targeting *in vitro* and *in vivo* of specific disease-related antigens for diagnosis and therapy in humans, there is a growing need for antibodies that combine the features of antigen specificity, low immunogenicity and high producibility.

[0006] Therefore, the problem underlying the present invention was to provide antibody proteins that combine the properties of specific binding to FAP, low immunogenicity in humans, and high producibility in recombinant systems.

Disclosure of the invention

[0007] The technical problem is solved by the embodiments characterized in the claims.

[0008] The present invention provides new antibody proteins having the complementary determining regions of the monoclonal antibody F19 (ATCC Accession No. HB 8269), said new antibody proteins specifically binding to fibroblast activation protein (FAP), characterised in that they have framework modifications resulting in the improved producibility in host cells as compared to a chimeric antibody having the variable regions of F19 and foreign constant regions.

[0009] As used herein, an "antibody protein" is a protein with the antigen binding specificity of a monoclonal antibody.

[0010] "Complementarity determining regions of a monoclonal antibody" are understood to be those amino acid

sequences involved in specific antigen binding according to Kabat (Kabat E. A., Wu T. T., Perry H. M., Gottesman K. S. and Foeller C. (1991) *Sequences of Proteins of Immunological Interest* (5th Edn). NIH Publication No. 91-3242. U.S. Department of Health and Human Services, Public Health Service, National Institutes of Health, Bethesda, MD.) in connection with Chothia and Lesk (Chothia and Lesk, J. Mol. Biol., 196:901-917 (1987)).

[0011] As used herein, the term "framework modifications" refers to the exchange, deletion or addition of single or multiple amino acids in the variable regions surrounding the individual complementarity determining regions. Framework modifications may have an impact on the immunogenicity, producibility or binding specificity of an antibody protein.

[0012] "Fibroblast activation protein (FAP)", also designated fibroblast activation protein alpha (FAP α), is a membrane-bound glycoprotein belonging to the serine protease gene family (WO 97/34927). No shed or secreted form of FAP is known.

[0013] FAP can be characterized by its binding to the monoclonal antibody F19 (F19 is obtainable from the hybridoma cell line with the accession No. HB 8269 deposited at the ATCC).

[0014] The term "fibroblast activation protein specific binding" of an antibody protein is defined herein by its ability to specifically recognise and stably bind FAP-expressing human cells. The binding specificity of the proteins of the invention can be determined by standard methods for the evaluation of binding specificity such as described in an exemplary fashion in example 6, 8 and example 12.

[0015] The term "chimeric antibody" refers to an antibody protein having the light and heavy chain variable regions as described in figures 17 and 18 and foreign constant regions. "Foreign constant regions" as defined herein are constant regions which are different from the constant regions of F19. For comparing an antibody protein of the invention to a chimeric antibody it is to be understood that such a chimeric antibody must contain the same constant regions as said antibody protein. For the purpose of demonstration and comparison alone the human constant heavy and light chains as described in Figures 19 to 22 are used in an exemplary fashion.

[0016] To provide the antibody proteins of the present invention, the nucleic acid sequences of the heavy and light chain genes of the murine antibody designated F19 were determined from RNA extracted from F19 hybridoma cells (ATCC Accession No. HB 8269).

[0017] In one embodiment the present invention relates to antibody proteins having the complementary determining regions of the monoclonal antibody F19 (ATCC Accession No. HB 8269), said new antibody proteins specifically binding to fibroblast activation protein (FAP), characterized in that they have framework modifications resulting in the improved producibility in host cells as compared to a chimeric antibody having the variable regions of F19 and foreign constant regions, wherein said antibody protein is derived from the murine antibody designated F19 (ATCC Accession No. HB 8269).

[0018] To generate humanised FAP-specific antibody proteins a chimeric antibody was constructed, having variable regions of the light and heavy chains of F19 and human light and heavy constant regions, respectively. The construction and production of chimeric mouse/human antibodies is well known (Boulianne et al. (1984), referenced above) and demonstrated in an exemplary fashion in examples 1 and 2.

[0019] Therefore, in a further embodiment the invention relates to antibody proteins according to the invention, characterised in that they have a variable light chain region and a variable heavy chain region, each joined to a human constant region.

[0020] In particular, the variable region of the light chain was joined to a human kappa constant region and the variable region of the heavy chain was joined to a human gamma-1 constant region. Other human constant regions for humanising light and heavy chains are also available to the expert. A human kappa and a human gamma-1 constant regions were used for demonstrating the invention in an exemplary fashion only.

[0021] Therefore, in one particular embodiment the antibody proteins of the invention contain a human kappa constant region.

[0022] Also, in another particular embodiment the antibody proteins of the invention contain a human gamma-1 constant region.

[0023] One particular "chimeric F19 antibody" protein (cF19) consists of the light and heavy chain variable and constant regions described in Figures 17 to 22. cF19 demonstrates specific binding and high avidity to the FAP antigen. As demonstrated in example 2, the expression of cF19 in COS cells is poor, ranging from about 10 to 60 ng/ml, which is at least 10 fold less than most antibodies.

[0024] In an attempt to increase expression levels of cF19, the leader sequence of the F19 V_L region was changed by substitution of Proline to Leucine at position -9.

[0025] This single change in amino acid in the leader sequence resulted in at least doubling the amount of chimeric antibody produced in COS cells. For the expression of this particular chimeric antibody in COS cells the following mutated leader sequence of the light chain: MDSQAQVLM LLLLVSGTCG, and the following leader sequence of the heavy chain: MGWSWVFLFLSGTAGVLS were used.

[0026] According to the invention the term "improved producibility" in host cells refers to the substantial improvement of expression levels and/or purified antibody yields when compared with the expression levels and/or antibody yields of

a chimeric antibody without framework modifications as defined above. Two particular but not limiting examples for demonstrating improved producibility are exemplified for the COS cell expression system (in examples 2 and 5) and for the CHO cell expression system (in example 10 and 11).

[0027] While the mutation of the leader sequence only lead to the doubling of the expression yield of the chimeric F19 antibody, a substantial improvement as defined herein refers to an improvement in expression level and/or purification yield of at least a factor of 10.

[0028] In a preferred embodiment, the invention refers to antibody proteins, characterised in that their expression levels in crude media samples as determined by ELISA and/or purified antibody yields exceed the expression levels and/or purification yields of the chimeric antibodies without framework modifications by at least a factor of 10.

[0029] In more preferred embodiment, the invention refers to antibody proteins, characterised in that their expression levels in crude media samples as determined by ELISA and/or purified antibody yields exceed the expression levels and/or purification yields of the chimeric antibodies without framework modifications by at least a factor of 20.

[0030] In a most preferred embodiment, antibody proteins, characterised in that their expression levels in crude media samples as determined by ELISA and/or purified antibody yields exceed the expression levels and/or purification yields of the chimeric antibodies without framework modifications by at least a factor of 100.

[0031] Improved producibility of the recombinant antibody proteins of the invention can be demonstrated for eucaryotic cells in general as shown for COS (cells derived from the kidney of an African green monkey) and CHO (Chinese hamster ovary derived cells) eucaryotic cells (see examples 5 and 11). In a further embodiment, the present invention relates to recombinant antibody proteins characterised in that they display improved producibility in eucaryotic cells.

[0032] In a preferred embodiment the present invention relates to antibody proteins, wherein said eucaryotic cell is a chinese hamster ovary cell (CHO cell).

[0033] It was unexpectedly found that certain framework modifications of the light chain variable regions determine the improved producibility of the antibody proteins of the invention. Three versions of reshaped light chain variable regions, designated version A, B, and C, as described in Figures 1 to 6, were prepared.

[0034] Light chain variable region versions A, B, and C demonstrate substantially improved producibility in CHO cells (see example 11). While light chain variable region versions A and C differ from light chain variable region version B by only two common amino acid residues they display an even further substantial improvement in producibility. There is at least another 10 fold difference in antibody secretion levels between the human reshaped F19 light chain version B and versions A or C. Reshaped human F19 light chain version A and B only differ in their amino acid sequences by two residues at positions 36 (Tyr to Phe mutation) and 87 (Tyr to Asp mutation) (nomenclature according to Kabat). This negative effect on the secretory capability of antibodies containing the light chain variable region version B could have been indirect if the Tyr to Asp and Tyr to Phe mutations, considered individually or together, merely caused improper folding of the protein. But this is unlikely to be the case since antigen binding assays show that immunoglobulins containing F19 light chain version B have similar avidities to those paired with F19 light chain version A or C, suggesting that they were not grossly misfolded.

[0035] Residue 87 in reshaped human F19 light chain version B seems particularly responsible for the reduction of secretion when compared to versions A and C.

[0036] In a preferred embodiment, the present invention relates to antibody proteins according to the invention, wherein the amino acid in Kabat position 87 of the light chain region is not asparagine.

[0037] In a more preferred embodiment, the invention relates to antibody proteins according to the invention, wherein the amino acid in Kabat position 87 of the light chain region is selected from aromatic or aliphatic amino acids.

[0038] In a most preferred embodiment, the present invention relates to antibody proteins according to the invention, wherein the aromatic amino acid in Kabat position 87 of the light chain region is a tyrosine or phenylalanine.

[0039] In a further embodiment, the present invention also pertains to antibody proteins according to the invention, wherein the amino acid in Kabat position 36 of the light chain region is selected from aromatic amino acids.

[0040] In a particular embodiment the invention relates to the specific antibody proteins that may be prepared from the individually disclosed reshaped variable regions of the light and heavy chains.

[0041] Especially light chain variable region versions A and C are particularly suitable to practice the invention because of their exceptionally high producibility, while retaining full FAP-binding specificity and achieving low immunogenicity. This holds especially true when compared to the chimeric antibody having the variable regions of F19 and the same constant regions but also when compared to light chain version B.

[0042] Therefore, in one embodiment the present invention relates to antibody proteins that contain the variable region of the light chain as set forth in SEQ ID NO: 2. In a further embodiment the invention also relates to antibody proteins, characterised in that the variable region of the light chain is encoded by a nucleotide sequence as set forth in SEQ ID NO: 1.

[0043] In one embodiment the present invention relates to antibody proteins that contain the variable region of the light chain as set forth in SEQ ID NO: 6.

[0044] In a further embodiment the invention also relates to antibody proteins characterised in that the variable region

of the light chain is encoded by a nucleotide sequence as set forth in SEQ ID NO: 5.

[0045] The present invention also discloses several different variable regions of the heavy chain that work particularly well with the variable regions of the light chain versions A and C in terms of improved producibility.

[0046] In one embodiment the invention relates to antibody proteins containing a variable region of the heavy chain as set forth in any one of SEQ ID NOs: 8, 10, 12, 14.

[0047] In another embodiment the invention relates to antibody proteins characterised in that the variable region of the heavy chain is encoded by a nucleotide sequence as set forth in any one of SEQ ID NOs: 7, 9, 11, 13.

[0048] In a very particular embodiment the invention relates to antibody proteins containing the variable region of the light chain as set forth in SEQ ID NO: 2 and the variable region of the heavy chain as set forth in SEQ ID NOs: 12.

[0049] In a further particular embodiment the invention relates to antibody proteins characterised in that the variable region of the light chain is encoded by a nucleotide sequence as set forth in SEQ ID NO: 1 and the variable region of the heavy chain is encoded by a nucleotide sequence as set forth in SEQ ID NO: 11.

[0050] In a further particular embodiment the invention relates to antibody proteins containing the variable region of the light chain as set forth in SEQ ID NO: 2 and the variable region of the heavy chain as set forth in SEQ ID NOs: 8.

[0051] In a further particular embodiment the invention relates to antibody proteins characterised in that the variable region of the light chain is encoded by a nucleotide sequence as set forth in SEQ ID NO: 1 and the variable region of the heavy chain is encoded by a nucleotide sequence as set forth in SEQ ID NO: 7.

[0052] In a further aspect, the present invention relates to nucleic acid molecules containing the coding information for the antibody proteins according to the invention as disclosed above. Preferably, a nucleic acid molecule according to the present invention is a nucleic acid molecule containing a nucleotide sequence selected from SEQ ID NOs: 1, 3, 5, 7, 9, 11, 13, or 15.

[0053] A further aspect of the present invention is a recombinant DNA vector containing the nucleotide sequence of any one of the above-mentioned nucleic acids, especially when said nucleotide sequence is operationally linked to an expression control sequence as in expression vectors. Preferred is a recombinant DNA vector, said vector being an expression vector.

[0054] A further aspect of the present invention is a host cell carrying a vector as described, especially an expression vector. Such a host cell can be a procaryotic or eucaryotic cell. Preferably, such a host cell is a eucaryotic cell, a yeast cell, or a mammalian cell. More preferably, said host cell is an CHO (Chinese hamster ovary) cell or a COS cell.

[0055] Accordingly, a still further aspect of the present invention is a method of producing antibody proteins according to the invention. Such a method comprises the steps of:

- (a) cultivating a host cell as described above under conditions where said antibody protein is expressed by said host cell, and
- (b) isolating said antibody protein.

[0056] Mammalian host cells, preferably CHO or COS cells are preferred. Host cells for producing the antibody proteins of the invention may be transfected with a single vector containing the expression units for both, the light and the heavy chain. In one particular embodiment the method of producing antibody proteins according to the invention pertains to host cells, wherein said host cells are cotransfected with two plasmids carrying the expression units for the light and heavy chains respectively.

[0057] The antibody proteins of the invention provide a highly specific tool for targeting therapeutic agents to the FAP antigen. Therefore, in a further aspect, the invention relates to antibody proteins according to the invention, wherein said antibody protein is conjugated to a therapeutic agent. Of the many therapeutic agents known in the art, therapeutic agents selected from the group consisting of radioisotopes, toxins, toxoids, inflammotogenic agents, enzymes, anti-sense molecules, peptides, cytokines, and chemotherapeutic agents are preferred.

[0058] Among the radioisotopes gamma, beta and alpha-emitting radioisotopes may be used as a therapeutic agent. β -emitting radioisotopes are preferred as therapeutic radioisotopes. ^{186}Re henium, ^{188}Re henium, ^{131}I odine and ^{90}Y trium have been proven to be particularly useful β -emitting isotopes to achieve localized irradiation and destruction of malignant tumor cells. Therefore, radioisotopes selected from the group consisting of ^{186}Re henium, ^{188}Re henium, ^{131}I odine and ^{90}Y trium are particularly preferred as therapeutic agents conjugated to the antibody proteins of the invention.

[0059] A further aspect of the present invention pertains to antibody proteins according to the invention, characterised in that they are labeled. Such an FAP-specific labeled antibody allows for the localisation and/or detection of the FAP antigen *in vitro* and/or *in vivo*. A label is defined as a marker that may be directly or indirectly detectable. An indirect marker is defined as a marker that cannot be detected by itself but needs a further directly detectable marker specific for the indirect marker. Preferred labels for practicing the invention are detectable markers. From the large variety of detectable markers, a detectable marker selected from the group consisting of enzymes, dyes, radioisotopes, and biotin is most preferred.

[0060] A further aspect of the present invention relates to antibody proteins according to the invention, characterised

in that they are conjugated to an imageable agent. A large variety of imageable agents, especially radioisotopes, are available from the state of the art. For practicing the invention gamma-emitting isotopes are more preferred. Most preferred is ^{125}I iodine.

[0061] One aspect of the present invention relates to pharmaceutical compositions containing an antibody protein according to the present invention as described above and a pharmaceutically acceptable carrier useful for treating tumors, wherein said tumors are associated with activated stromal fibroblasts. There are two possible effector principles for an anti-tumor stroma immunotherapy that may act synergistically: (a) An unmodified (unconjugated, 'naked') antibody according to the invention may induce immune destruction or inflammatory reactions in the tumor stroma while (b) an antibody conjugated to a therapeutic agent, such as for example, a radioisotope or other toxic substance, may achieve localized irradiation and destruction of the malignant tumor cells.

[0062] One further embodiment are pharmaceutical compositions containing an antibody protein according to the invention conjugated to a therapeutic agent as described above and a pharmaceutically acceptable carrier useful for treating tumors, wherein said tumors are associated with activated stromal fibroblasts. Another embodiment pertains to pharmaceutical compositions containing an antibody protein according to the present invention conjugated to an imageable agent as described above and a pharmaceutically acceptable carrier useful for imaging the presence of activated stromal fibroblasts in a healing wound, inflamed skin or a tumor, in a human patient. A most preferred embodiment relates to the pharmaceutical compositions mentioned above, wherein said tumors are tumors selected from the cancer group consisting of colorectal cancers, non-small cell lung cancers, breast cancers, head and neck cancer, ovarian cancers, lung cancers, invasive bladder cancers, pancreatic cancers and cancers metastatic of the brain.

[0063] In an animal or human body, it can prove advantageous to apply the pharmaceutical compositions as described above via an intravenous or other route, e.g. systemically, locally or topically to the tissue or organ of interest, depending on the type and origin of the disease or problem treated, e.g. a tumor. For example, a systemic mode of action is desired when different organs or organ systems are in need of treatment as in e.g. systemic autoimmune diseases, or allergies, or transplantations of foreign organs or tissues, or tumors that are diffuse or difficult to localise. A local mode of action would be considered when only local manifestations of neoplastic or immunologic action are expected, such as, for example local tumors.

[0064] The antibody proteins of the present invention may be applied by different routes of application known to the expert, notably intravenous injection or direct injection into target tissues. For systemic application, the intravenous, intravascular, intramuscular, intraarterial, intraperitoneal, oral, or intrathecal route are preferred.

[0065] A more local application can be effected subcutaneously, intracutaneously, intracardially, intralobally, intramedullarily, intrapulmonarily or directly in or near the tissue to be treated (connective-, bone-, muscle-, nerve-, epithelial tissue). Depending on the desired duration and effectiveness of the treatment, pharmaceutical antibody compositions may be administered once or several times, also intermittently, for instance on a daily basis for several days, weeks or months and in different dosages.

[0066] For preparing suitable antibody preparations for the applications described above, the expert may use known injectable, physiologically acceptable sterile solutions. For preparing a ready-to-use solution for parenteral injection or infusion, aqueous isotonic solutions, such as e.g. saline or corresponding plasmaprotein solutions are readily available. The pharmaceutical compositions may be present as lyophilisates or dry preparations, which can be reconstituted with a known injectable solution directly before use under sterile conditions, e.g. as a kit of parts. The final preparation of the antibody compositions of the present invention are prepared for injection, infusion or perfusion by mixing purified antibodies according to the invention with a sterile physiologically acceptable solution, that may be supplemented with known carrier substances or/and additives (e.g. serum albumine, dextrose, sodium bisulfite, EDTA).

[0067] The amount of the antibody applied depends on the nature of the disease.

[0068] Furthermore, one aspect of the present invention relates to the use of the antibody proteins according to the invention for the treatment of cancer. In a preferred embodiment the present invention relates to the use of antibody proteins according to the invention conjugated to a therapeutic agent as described above for the treatment of cancer. In another preferred embodiment the present invention relates to the use of antibody proteins according to the invention conjugated to an imageable agent for imaging activated stromal fibroblasts. In a further preferred embodiment the present invention relates to the use of labeled antibody proteins according to the invention for detecting the presence of activated stromal fibroblasts in a sample.

[0069] One aspect of the invention relates to a method of treating tumors, wherein the tumor is associated with activated stromal fibroblasts capable of specifically forming a complex with antibody proteins according to the invention, present as naked/unmodified antibodies, modified antibody proteins, such as e.g. fusion proteins, or antibody proteins conjugated to a therapeutic agent, which comprises contacting the tumor with an effective amount of said antibodies. In a preferred embodiment the present invention relates to a method of treating tumors as mentioned above, wherein the tumor is a tumor having cancer cells selected from the cancer group consisting of colorectal cancers, non-small cell lung cancers, breast cancers, head and neck cancer, ovarian cancers, lung cancers, invasive bladder cancers, pancreatic cancers and metastatic cancers of the brain. The method of treating tumors as described above may be effected in

in vitro or *in vivo*.

[0070] A further aspect of the invention relates to a method of detecting the presence of activated stromal fibroblasts in wound healing, inflammation or in tumors, characterised in that

- 5 (a) a sample, possibly containing activated stromal fibroblasts, is contacted with an antibody protein according to the invention under conditions suitable for the formation of a complex between said antibody and antigen,
- (b) detecting the presence of said complex, thereby detecting the presence of activated stromal fibroblasts in wound healing, inflammation or a tumor.

10 **[0071]** In a preferred embodiment, the present invention relates to a method of detecting the presence of activated stromal fibroblasts in a tumor, wherein the tumor is a tumor having cancer cells selected from the cancer group consisting of colorectal cancers, non-small cell lung cancers, breast cancers, head and neck cancer, ovarian cancers, lung cancers, bladder cancers, pancreatic cancers and metastatic cancers of the brain. Most preferred antibody proteins of the invention are those which are characterised in that they are labeled as mentioned above.

15 **[0072]** A further aspect of the invention relates to a method of imaging the presence of activated stromal fibroblasts in a healing wound, inflamed skin or a tumor, in a human patient, characterised in that

- (a) an antibody protein according to the present invention conjugated to an imageable agent is administered to a human patient under conditions suitable for the formation of an antibody-antigen complex,
- 20 (b) imaging any complex formed in this manner,
- (c) thereby imaging the presence of activated stromal fibroblasts in a human patient.

[0073] In a preferred embodiment the present invention relates to a method of imaging the presence of activated stromal fibroblasts as described above in tumors, wherein the tumor is a tumor having cancer cells selected from the cancer group consisting of colorectal cancers, non-small cell lung cancers, breast cancers, head and neck cancer, ovarian cancers, lung cancers, bladder cancers, pancreatic cancers and metastatic cancers of the brain.

[0074] In a further aspect the present invention relates to a method of detecting tumor-stroma, characterised in that

- 30 (a) a suitable sample is contacted with an antibody protein according to the present invention, under conditions suitable for the formation of an antibody-antigen complex,
- (b) detecting the presence of any complex so formed,
- (c) relating the presence of said complex to the presence of tumor-stroma.

[0075] Antibody proteins for practicing the invention are preferably labelled with a detectable marker.

35 **[0076]** In a further aspect the present invention relates to a method of imaging tumor-stroma in a human patient, which comprises

- (a) administering to the patient an antibody according to the invention conjugated to an imageable agent as described above under conditions suitable for the formation of an antibody-antigen complex,
- 40 (b) imaging any complex so formed, and thereby imaging the presence of tumor-stroma in a human patient.

Figure legends

[0077]

45

Fig. 1. DNA sequence of F19 human reshaped light chain variable region version A (hF19L_A) SEQ ID NO:1.

Fig. 2. Amino acid sequence of F19 human reshaped light chain variable region version A (hF19L_A) SEQ ID NO: 2.

50 **Fig. 3.** DNA sequence of F19 human reshaped light chain variable region version B (hF19L_B) SEQ ID NO: 3. Nucleotides differing from version A are underlined and in bold type.

Fig. 4. Amino acid sequence of F19 human reshaped light chain variable region version B (hF19L_B) SEQ ID NO: 4. Amino acids differing from version A are underlined and in bold type.

55

Fig. 5. DNA sequence of F19 human reshaped light chain variable region version C (hF19L_C) SEQ ID NO:5. Nucleotides differing from version A are underlined and in bold type.

Fig. 6. Amino acid sequence of F19 human reshaped light chain variable region version C (hF19L_C) SEQ ID NO: 6. Amino acids differing from version A are underlined and in bold type.

Fig. 7. DNA sequence of F19 human reshaped variable region heavy chain version A (hF19H_A) SEQ ID NO: 7.

Fig. 8. Amino acid sequence of F19 human reshaped heavy chain variable region version A (hF19H_A) SEQ ID NO: 8

Fig. 9. DNA sequence of F19 human reshaped heavy chain variable region version B (hF19H_B) SEQ ID NO: 9. Nucleotides differing from version A are underlined and in bold type.

Fig. 10. Amino acid sequence of F19 human reshaped heavy chain variable region version B (hF19H_B) SEQ ID NO: 10. Amino acids differing from version A are underlined and in bold type.

Fig. 11. DNA sequence of F19 human reshaped heavy chain variable region version C (hF19H_C) SEQ ID NO: 11. Nucleotides differing from version A are underlined and in bold type.

Fig. 12. Amino acid sequence of F19 human reshaped heavy chain variable region version C (hF19H_C) SEQ ID NO: 12. Amino acids differing from version A are underlined and in bold type.

Fig. 13. DNA sequence of F19 human reshaped heavy chain variable region version D (hF19H_D) SEQ ID NO: 13. Nucleotides differing from version A are underlined and in bold type.

Fig. 14. Amino acid sequence of F19 human reshaped heavy chain variable region version D (hF19H_D) SEQ ID NO: 14. Amino acids differing from version A are underlined and in bold type.

Fig. 15. DNA sequence of F19 human reshaped heavy chain variable region version E (hF19H_E) SEQ ID NO: 15. Nucleotides differing from version A are underlined and in bold type.

Fig. 16. Amino acid sequence of F19 human reshaped heavy chain variable region version E (hF19H_E) SEQ ID NO: 16. Amino acids differing from version A are underlined and in bold type

Fig. 17. Amino acid sequence of F19 chimeric light chain variable region (chF19LC) SEQ ID NO: 17.

Fig. 18. Amino acid sequence of F19 chimeric heavy chain variable region (chF19HC) SEQ ID NO: 18.

Fig. 19. DNA sequence of human kappa light constant chain SEQ ID NO: 19.

Fig. 20. Amino acid sequence of human light constant chain SEQ ID NO: 20.

Fig. 21. DNA sequence of human heavy constant chain SEQ ID NO: 21.

Fig. 22. Amino acid sequence of human heavy constant chain SEQ ID NO: 22.

Fig. 23. Mammalian cell expression vectors used to produce chimeric and reshaped human antibodies with human kappa light chains and human gamma-1 heavy chains.

A. Light chain expression vector: pKN100

B. Heavy chain expression vector: pG1D105

Fig 24. DNA and amino acid sequences of mouse F19 light chain variable region as modified for use in the construction of chimeric F19 light chain. Restriction sites are indicated by bold letters. The Kozak sequence, CDR's 1 to 3 and the splice donor site are underlined.

Fig 25. DNA and amino acid sequences of mouse F19 heavy chain variable region as modified for use in the construction of chimeric F19 heavy chain. Restriction sites are indicated by bold letters. The Kozak sequence and the splice donor site are underlined.

Fig. 26. DNA sequence of F19 chimeric antibody cloned into pKN100 mammalian expression vector. Restriction sites are indicated by bold letters and underlined. CDR's 1 to 3 and the splice donor site are underlined. This is the DNA sequence of the mouse F19 light chain inside the pKN100 eukaryotic expression vector. This vector has a cDNA version of the human kappa constant region gene (allotype Km(3)) terminated by a strong artificial termination sequence. In addition, the Neo selection gene is also terminated by this artificial sequence and is also in the same orientation as the kappa light chain expression cassette.

The essential components of the pKN100 eukaryotic expression vector are:

1 - 6	= EcoRI site
7 - 1571	= HCMVi promoter/enhancer
583 - 587	= TATAA box
610	= Start of transcription
728 - 736	= Splice donor site
731	= Beginning of intron
1557	= End of intron
1544 - 1558	= Splice acceptor site
1590 - 1598	= Kozak sequence
1599 - 1658	= peptide leader sequence
1659 - 1997	= mouse F19 light chain
1996 - 2004	= splice donor site
2011 - 2657	= cDNA copy of human Kappa constant region (Km(3)) gene
2664 - 2880	= Artificial spaC2 termination sequence
2887 - 7845	= This is the pSV2neo vector DNA fragment comprising of the Amp-resistance gene (in the opposite orientation), the ColEI and SV40 origins of replication and the Neo-resistance gene (in the same orientation as the HCMVi-KCT cassette)
7852 - 8068	= Artificial spaC2 termination signal

This sequence ends immediately upstream of the EcoRI site (position 1-6) at the beginning of the sequence. As a vector this DNA sequence would be circular.

Fig. 27. DNA sequence of F19 chimeric antibody cloned into pg1d105 mammalian expression vector. Restriction sites are indicated by bold letters and underlined. CDR's 1 to 3 and the splice donor site are underlined. This is the DNA sequence of the eukaryotic expression vector pG1D105 containing the mouse F19 heavy chain variable region. This vector contains a cDNA version of the human gamma-1 constant region (allotype G1m^{Non-a}).

The essential components of the construct are:

1 - 2501	= pBR322 based sequence including Ampicillin resistance gene and ColEI origin plus the SV40 origin and the crippled SV40 early promoter
2502 - 3226	= dhfr gene
3233 - 4073	= SV40 poly A sequence etc.
4074 - 4079	= ligated BamHI and BglII site (BstYI)
4080 - 4302	= SPA site plus C2 termination signal
4303 - 5867	= HCMVi promoter
5879 - 5885	= unique HindIII restriction site for cloning of immunoglobulin variable genes
5886 - 5894	= Kozak sequence
5895 - 5951	= signal peptide
5952 - 6323	= mouse F19 heavy chain
6323 - 6330	= splice donor site
6331 - 6336	= unique BamHI restriction site for cloning of immunoglobulin variable genes
6337 - 7388	= cDNA copy of human gamma-1 constant regions preceded by a 62 bp intron
7389 - 7709	= Arnie termination sequence

The human gamma-1 constant region used in this construct has a G1m^{Non-a} allotype which is defined by a Glutamic acid (E) residue at position 356 (according to Eu numbering) and a Methionine (M) residue at position 358 (according to Eu numbering). These two residues are underlined in the sequence above.

Fig. 28. PCR-based method for the construction of human reshaped F19 light chain. This figure provides a schematic overview of the strategy of construction. The dotted lines indicate a complementary sequence of at least 21

bases between the primers.

Fig. 29. Nucleotide and deduced amino acid sequences of reshaped human F19 light chain variable regions version A, B and C. Nucleotide and deduced amino acid sequences are aligned and compared with that of version A, dashes indicate nucleotide identity, dots indicate amino acid identity with this sequence. Amino acids are numbered according to Kabat *et al.* (1991). The locations of CDRs are indicated in boxes.

Fig. 30. DNA sequence of F19 L_A (human reshaped light chain version A) cloned into pKN100 mammalian expression vector. Restriction sites are indicated by bold letters and underlined. CDR's 1 to 3 and the splice donor site are underlined. This is the DNA sequence of the reshaped F19 light chain version A cloned into pKN100 eukaryotic expression vector. This vector has a cDNA version of the human kappa constant region gene (allotype Km(3)) terminated by a strong artificial termination sequence. In addition, the Neo selection gene is also terminated by this artificial sequence and is also in the same orientation as the kappa light chain expression cassette.

The components of the vector are:

7 - 1571	= HCMVi promoter/enhancer
583 - 587	= TATAA box.
610	= Start of transcription.
728 - 736	= Splice donor site.
731	= Beginning of intron.
1557	= End of intron.
1544 - 1558	= Splice acceptor site.
1590 - 1598	= Kozak sequence
1599 - 1658	= peptide leader sequence
1659 - 1997	= reshaped F19 light chain version A
1996 - 2004	= splice donor site
2011 - 2657	= cDNA copy of human kappa constant region (Km(3)) gene.
2664 - 2880	= Artificial spaC2 termination sequence.
2887 - 7845	= This is the pSV2neo vector DNA fragment comprising of the Amp-resistance gene (in the opposite orientation), the ColEI and SV40 origins of replication and the Neo-resistance gene (in the same orientation as the HCMVi-KCT cassette).
7852 - 8068	= Artificial spaC2 termination signal.

This sequence ends immediately upstream of the EcoRI site (position 1-6) at the beginning of the sequence below. As a vector this DNA sequence would be circular.

Fig. 31. PCR-based method for the construction of human reshaped F19 heavy chain. This figure provides a schematic overview of the strategy of construction. The dotted lines indicate a complementary sequence of at least 21 bases between the primers.

Fig. 32. Nucleotide and deduced amino acid sequences of reshaped human F19 heavy chain variable region versions a to e. Nucleotide and deduced amino acid sequences are aligned and compared with that of version A, dashes indicate nucleotide identity, dots indicate amino acid identity with this sequence. Amino acids are numbered according to Kabat *et al.* (1991). The location of CDRs is indicated by boxes.

Fig. 33. DNA sequence of F19Ha (human reshaped heavy chain version a) cloned into pg1d105 mammalian expression vector. Restriction sites are indicated by bold letters and underlined. CDR's 1 to 3 and the splice donor site are underlined. This is the DNA sequence of the eukaryotic expression vector pG1D105 containing the reshaped version A of F19 heavy chain variable region. This vector contains a cDNA version of the human gamma-1 constant region (allotype G1m^{Non-a}).

The essential components of the construct are:

1 - 2501	= pBR322 based sequence including Ampicillin resistance gene and ColEI origin plus the SV40 origin and the crippled SV40 early promoter
2502 - 3226	= dhfr gene
3233 - 4073	= SV40 poly A sequence etc.
4080 - 4302	= SPA site plus C2 termination signal
4303 - 5867	= HCMVi promoter/enhancer

5879 - 5885 = unique HindIII restriction site for cloning of immunoglobulin variable genes
 5886 - 5894 = Kozak sequence
 5895 - 5951 = signal peptide
 5952 - 6323 = reshaped F19 heavy chain version A
 6323 - 6330 = splice donor site
 6331 - 6336 = unique BamHI restriction site for cloning of immunoglobulin variable genes
 6337 - 7388 = cDNA copy of human gamma-1 constant regions preceded by a 62 bp intron
 7389 - 7709 = Arnie termination sequence

The human gamma-1 constant region used in this construct has a G1m^{Non-a} allotype which is defined by a Glutamic acid (E) residue at position 356 (according to Eu numbering) and a Methionine (M) residue at position 358 (according to Eu numbering). These two residues are underlined in the sequence above.

Fig. 34. Heavy (panel A) and light (panel B) chains RNA splicing events taking place during antibody F19 expression in mammalian cells - schematic overview.

A. Heavy chain RNA splicing

B. Kappa light chain RNA splicing

Fig. 35. Concentration dependence of L_AH_C supernatant binding to CD8-FAP.

Fig. 36. Binding of biotinylated L_AH_C to human FAP.

Fig. 37. CD8-FAP carries the F19 epitope as detected with cF19.

Examples

Example 1: Construction of mouse - human chimeric genes

[0078] The chimeric F19 (cF19) antibody was designed to have the mouse F19 V_L and V_H regions linked to human kappa and gamma-1 constant regions, respectively. PCR primers were used to modify the 5'- and 3'- sequences flanking the cDNA sequences coding for the mouse F19 V_L and V_H regions (Table 1). PCR primers specific for F19 light chain V-region were designed. These adapted mouse F19 variable regions were then subcloned into mammalian cell expression vectors already containing the human kappa (pKN100 vector) or gamma-1 (pG1D105 vector) constant regions (Figure 23).

[0079] These vectors employ the human cytomegalovirus (HCMV) promoter/enhancer to efficiently transcribe the light and heavy chains. The vectors also contain the SV40 origin of replication to permit efficient DNA replication and subsequent protein expression in cos cells. The expression vectors were designed to have the variable regions inserted as HindIII-BamHI DNA fragments. PCR primers were designed to introduce these restrictions sites at the 5'- (HindIII) and 3'- (BamHI) ends of the cDNAs coding for the V-regions. In addition the PCR primers were designed to introduce the Kozak sequence (GCCGCCACC) at the 5'-ends of both the light and heavy chain cDNAs to allow efficient translation (Kozak M.: At least six nucleotides preceding the AUG initiator codon enhance translation in mammalian cells. *J. Mol. Biol.* (1987) **196**: 947), and to introduce splice donor sites at the 3'-ends of both the light and heavy chain cDNAs for the variable regions to be spliced to the constant regions. The PCR primers used in the construction of the chimeric F19 light and heavy chains are shown in Table 1. The DNA and amino acid sequences of the mouse F19 V_L and V_H regions as adapted for use in the construction of chimeric F19 light and heavy chains are shown in Figures 24 and 25. The DNA sequences of mouse F19 light and heavy chains cloned into the eukaryotic expression vectors pKN100 and pG1D105, respectively, are shown in Figures 26 and 27.

TABLE 1: PCR primers for the construction of chimeric F19 antibody.**A. Light chain variable region**

1. Primer for the construction of the 5'-end (37mer)

5' CAGA **AAGCTT** GCCGCCACC ATG GAT TCA CAG GCC CAG 3'**HindIII** Kozak sequence M D S Q A Q

2. Primer for the construction of the 3'-end (35mer)

5' CCGA **GGATCC** ACTCACG TTT CAG CTC CAG CTT GGT 3'**BamHI** Splice donor site**B. Heavy chain variable region**

1. Primer for the construction of the 5'-end (37mer)

5' CAGA **AAGCTT** GCCGCCACC ATG GGA TGG AGC TGG GTC 3'**HindIII** Kozak sequence M G W S W V

2. Primer for the construction of the 3'-end (35mer)

5' CCGA **GGATCC** ACTCACC TGA GGA GAC GGT GAC TGA 3'**BamHI** Splice donor site**Example 2: Expression and binding activity of chimeric F19 antibody**

[0080] The two plasmid DNAs coding for the chimeric F19 light and heavy chains (see example 1) were co-transfected into cos cells to look for transient expression of chimeric F19 antibody as described below. After 72 h incubation, the medium was collected, centrifuged to remove cellular debris, and analysed by ELISA for the production of a human IgG1-like antibody. The cos cell supernatant containing the chimeric F19 antibody was analysed for its ability to bind to HT 1080 cells (see example 13) expressing the FAP antigen on their surface.

Transfection of cos cells using electroporation

[0081] The mammalian expression vectors pg1d105 and pKN100 containing the chimeric or reshaped human heavy and light chains versions, respectively, were tested in cos cells to look for transient expression of F19 antibodies. Cos

7 cells were passaged routinely in DMEM (Gibco BRL cat. #41966) containing penicillin (50 IU/ml), streptomycin (50 µg/ml), L-glutamine and 10% heat-inactivated gamma globulin-free foetal calf serum (FCS, Harlan Sera-Lab cat. # D0001). The DNA was introduced into the cos cells by electroporation using the Gene Pulsar apparatus (BioRad). DNA (10 µg of each vector) was added to a 0.8 ml aliquot of 1×10^7 cells/ml in Phosphate-buffered saline (PBS, Ca^{2+} and Mg^{2+} free). A pulse was delivered at 1,900 volts, 25 µF capacitance. After a 10 min recovery period at ambient temperature the electroporated cells were added to 8 ml of DMEM containing 5% FCS. After 72h incubation at 37°C, the medium was collected, centrifuged to remove cellular debris, and stored under sterile conditions at 4°C for short periods of time, or at -20°C for longer periods.

ELISA method for measuring assembled IgG1/kappa antibody concentrations in cos cell supernatants

[0082] Samples of antibodies produced in transfected cos cells were assayed by ELISA to determine how much reshaped human antibody had been produced. For the detection of human antibody, plates were coated with goat anti-human IgG (Fcγ fragment specific) antibody (Jackson ImmunoResearch Laboratories Inc., #109-005-098). The samples from cos cells were serially diluted and added to each well. After incubation for 1h at 37°C and washing, horseradish peroxidase conjugated goat anti-human kappa light chain (Sigma, A-7164) was added. After incubation for 30 mins at 37°C and washing, K-blue substrate (mixer of 3,3',5,5' tetramethylbenzidine and hydrogen peroxide, Bionostics Limited, #KB175) was added. After standing at room temperature for 30 mins, the reaction was stopped using Red Stop solution (Bionostics Limited, #RS20) and the optical density read on a microplate reader at 650 nm. Purified human IgG1/Kappa antibody (Sigma, I-3889) of known concentration was used as a standard.

[0083] The expression of chimeric F19 antibody in COS cells was poor (Table 2), between 10 and 60 ng/ml which is at least 10 fold less than most antibodies.

[0084] In an attempt to increase expression levels of the chimeric F19 antibody, the leader sequence of F19 V_L region was changed by substitution of Leucine to Proline at position -9. This single change in amino acid in the leader sequence resulted in at least doubling the amount of chimeric antibody produced in COS cells.

[0085] The test results show that chimeric F19 binds specifically and with the expected avidity to the FAP target.

TABLE 2

Chimeric F19 antibody concentrations in COS cell supernatants (These are the results of three independent transfections)		
Transfected Antibody components		Human γ1/K
Heavy chain	Kappa light chain	[in µg/ml]
cF19	cF19 (F19 leader sequence)	0.060
cF19	cF19 (mutated leader sequence)	0.212
cF19	cF19 (F19 leader sequence)	0.056
cF19	cF19 (mutated leader sequence)	0.108
cF19	cF19 (F19 leader sequence)	0.011
cF19	cF19 (mutated leader sequence)	0.087

Example 3: Construction of the reshaped human F19 light chain versions a to c (La-Lb)

[0086] The construction of the first version of reshaped human F19 V_L region (La) was carried out using overlapping PCR fragments in a method similar to that described by Daugherty B. L., DeMartino J. A., Law M. F., Kawka D. W., Singer I. I. and Mark G. E. (1991) Polymerase chain reaction facilitates the cloning, CDR-grafting, and rapid expression of a murine monoclonal antibody directed against the CD18 component of leukocyte integrins. *Nucl. Acids Res.* **19**: 2471. Ten oligonucleotides were synthesised that consisted of five primer pairs, APCR1-vla1, vla2-vla3, vla4-vla5, vla6-vla7, and vla8-APCR4 (Table 3 and Figure 28). There was an overlapping sequence of at least 21 bases between adjacent pairs (Figure 28). APCR1 and APCR4 hybridised to the flanking pUC19 vector sequences. The mutagenic primers were designed such that their 5' end immediately followed the wobble position of a codon. This strategy was used to counteract the gratuitous addition of one nucleotide to the 3' end of the strand complementary to the mutagenic primer by the DNA polymerase during PCR (Sharrocks A. D. and Shaw P. E. (1992) Improved primer design for PCR-based, site-directed mutagenesis. *Nucl. Acids Res.* **20**: 1147). The appropriate primer pairs (0.2 µM of each) were combined

with 10ng of version "b" of reshaped human L25V_L region cDNA, and 1 unit of AmpliTaq (Perkin Elmer Cetus) DNA polymerase in 50μl of PCR buffer containing 10mM Tris-HCl (pH8.3), 50mM KCl, 200μM dNTPs, and 1.5mM MgCl₂. This was overlaid with mineral oil and PCR was performed for 25 cycles, each cycle consisting of a denaturation step at 94°C for 1 min, a primer annealing step at 55°C for 1 min, and an extension step at 72°C for 2 mins. This was followed by a single cycle consisting of a further elongation step at 72°C for 10 mins followed by cooling to 4°C. The ramp time between the primer-annealing and extension steps was 2.5 mins. The PCR products of the five reactions (A, B, C, D and E) were then purified by gel electrophoresis followed by DNA elution using Wizard PCR preps (Promega). PCR products A, B, C, D, and E were assembled by their complementarity to one another. In the second set of PCR reactions, PCR products B and C, and D and E, (50ng of each) were added to 50μl PCR reactions (as described above) each containing 1 unit of AmpliTaq (Perkin Elmer Cetus) DNA polymerase. The reactions were cycled for 20 cycles as described above with the exception that the annealing temperature was raised to 60°C. In the third set of PCR reactions, PCR products F and G were PCR-amplified using 1 μl of each prior PCR reaction and the appropriate pair of PCR primers (vla2-vla5 or vla6-APCR4). The PCR reactions contained 1 unit of AmpliTaq DNA polymerase in 50 μl PCR reaction (as described above) and were amplified for 25 cycles as in the first stage. In the fourth set of PCR reactions, the PCR product H was PCR-amplified using 1 μl of each prior PCR reaction and the vla2-APCR4 pair of PCR primers. Finally, PCR products A and H were assembled by their own complementarity in a two step-PCR reaction similar to that described above using RSP and UP as the terminal primers. The fully assembled fragment representing the entire reshaped human F19 V_L region including a leader sequence was digested with HindIII and BamHI and cloned into pUC19 for sequencing. A clone having the correct DNA sequence was designated reshF19La (Figure 29) and was then subcloned into the eukaryotic expression vector pKN100. The DNA sequence of reshF19La cloned into pKN100 is shown in Figure 30.

[0087] The second version of reshaped human F19 V_L region (Lb) was constructed using the same scheme as that described for La but where vla4 and vla7 primers were substituted by vlb4 and vlb7 respectively (Table 3). The DNA sequence of Lb is shown in Figure 29.

[0088] The third version of reshaped human F19 V_L region (Lc) was constructed using the QuikChange™ site-directed mutagenesis kit from Stratagene. The QuikChange site-directed mutagenesis method was performed according to the manufacturer's instructions, using reshF19La in pKN100 vector as double stranded DNA template. The mutagenic oligonucleotide primers F19Lc-sense and F19Lc-antisense (Table 3) for use in this protocol were designed according to the manufacturers instructions. Briefly, both the mutagenic primers contained the desired point mutation (codon TTT at Kabat residue position 49 (Phe) changed to TAT coding for Tyr) and annealed to the same sequence on opposite strands of La in pKN100 vector. The point mutation was verified by DNA sequencing the entire V_L region. The DNA sequence of Lc is shown in Figure 29. To eliminate the possibility that random mutations occurred in the pKN100 during the PCR reaction, the V_L region was cut out of the pKN100 vector as an HindIII/BamHI fragment and re-subcloned into an unmodified pKN100 vector cut with the same two restriction enzymes beforehand.

TABLE 3: PCR primers for the construction of reshaped human F19 light chain variable regions

1. Primers for the synthesis of version "a"

F19vla1 (36 mer):

5' GTCATCACAATGTCTCCGGAGGAACCTGGAACCCAG 3'

F19vla2 (29 mer):

5' CTCCGGAGACATTGTGATGACCCAATCTC 3'

F19vla3 (45 mer):

5' GAATATAAAAGGCTCTGACTGGACTTGCAGTTGATGGTGGCCCTC 3'

F19vla4 (72 mer):

5' CAGTCAGAGCCTTTTATATTCTAGAAATCAAAGAACTACTTGGCCTGGTAT
CAGCAGAAACCAGGACAGCC 3'

F19vla5 (44 mer):

5' ACCCCAGATTCCCTAGTGCTAGCCCCAAAGATGAGGAGTTTGGG 3'

F19vla6 (67 mer):

5' TAGCACTAGGGAATCTGGGGTACCTGATAGGTTTCAGTGGCAGTGGGTTTG
GGACAGACTTCACCCTC 3'

F19vla7 (53 mer):

5' GTCCCTTGTCCGAACGTGAGCGGATAGCTAAAATATTGCTGACAGTAA
TAAAC 3'

F19vla8 (33 mer):

5' GCTCACGTTCGGACAAGGGACCAAGGTGGAAAT 3'

2. Primers for the synthesis of version "b"

F19vIb4 (72 mer):

5' CAGTCAGAGCCTTTTATATTCTAGAAATCAAAGAACTACTTGGCCTGG
TTCCAGCAGAAACCAGGACAGCC 3'

F19vIb7 (57 mer):

5' GTCCCTTGTCCGAACGTGAGCGGATAGCTAAAATATTGCTGACAGTCATA
AACTGCC 3'

3. Primers for the synthesis of version "c"

F19Lc-sense (34 mer):

5' CCCAAACTCCTCATCTATTGGGCTAGCACTAGGG 3'

F19Lc-antisense (34 mer):

5' CCCTAGTGCTAGCCCAATAGATGAGGAGTTTGGG 3'

4. Primers hybridizing to the flanking PUC19 vector sequences

APCR1 (17 mer, sense primer): 5' TACGCAAACCGCCTCTC 3'

APCR4 (18 mer, anti-sense primer): 5' GAGTGCACCATATGCGGT 3'

RSP (-24) (16 mer, sense primer): 5' AACAGCTATGACCATG 3'

UP (-40) (17 mer, anti-sense primer): 5' GTTTTCCCAGTCACGAC 3'

Example 4: Construction of the reshaped human F19 heavy chain versions a to e (Ha-He)

[0089] Version "a" of reshaped human F19 V_H regions (Ha) was constructed using the same PCR methods as described for the construction of version "a" of reshaped human F19 V_L region (La) (Figure 31). The template DNA was version "a" of reshaped human 226 V_H (Léger O. J. P., Yednock T. A., Tanner L., Horner H. C., Hines D. K., Keen S., Saldanha J., Jones T., Fritz L. C. and Bendig M. M. (1997). Humanization of a mouse antibody against human alpha-4 integrin: a potential therapeutic for the treatment of multiple sclerosis. *Hum. Antibod.* **8**: 3). Six PCR primers were designed and synthesized for the construction of version "a" of reshaped human F19 V_H region (Table 4). PCR products A, B, C, and D were obtained using APCR1-Vha1, Vha2-Vha3, Vha4-Vha5 and Vha6-APCR4 as PCR primer pairs, respectively. The PCR conditions were essentially as described for the construction of reshaped human F19 V_L region. A clone having the correct DNA sequence was designated reshF19Ha (Figure 32) and was then subcloned into the eukaryotic expression vector pG1D105. The DNA sequence of reshF19Ha cloned into pG1D105 is shown in Figure 33.

[0090] The third version of reshaped human F19 V_H region (Hc) was constructed using the same scheme as that described for Ha but where Vha4 primer was substituted by Vhc4 (Table 4). The DNA sequence of Hc is shown in Figure 32. The second (Hb) and fourth (Hd) version of reshaped human F19 V_H region were constructed based on the PCR-mutagenesis methods of Kamman et al. (Kamman M., Laufs J., Schell J. and Gronenborn B. (1989) Rapid insertional mutagenesis of DNA by polymerase chain reaction (PCR). *Nucl. Acids Res.* **17**: 5404). For Hb and Hd, a mutagenic primer F19VHbd6 (Tyr-91 to Phe-91, Table 4) was used paired with APCR4 in PCR reactions with Ha and Hc as the template DNA, respectively. The PCR products VHb and VHd were restriction enzyme digested with PstI and BamHI and subcloned into reshF19Ha and reshF19Hc, respectively, previously digested with the same two restriction enzymes. The DNA sequences of Hb and Hd are shown in Figure 32.

[0091] Version e of reshaped human F19 V_H region (He) was constructed based on the PCR-mutagenesis methods of Kamman et al. (1989) already mentioned above:

[0092] For reshF19He mutagenic primer F19MscIHe (Table 5) was used paired with primer F19V_HHindIII (Table 5) in PCR reactions with Hc cloned in pg1d105 mammalian expression vector as the template DNA. The appropriate primer pairs (0.2μM of each) were combined with 10ng of cDNA of version "a" of reshaped human 226 V_H region in 100μl of PCR buffer containing 10mM KCl, 10mM (NH₄)₂SO₄, 20mM Tris-HCl (pH 8.8) 2mM MgSO₄, 0.1% Triton X-100 and 200μM dNTPs. Reaction mixtures were overlaid with mineral oil and kept at 94°C for 5 mins. Then 1 unit of Deep Vent DNA polymerase (New England Biolabs) was added ("Hot Start" PCR; Chou Q., Russell M., Birch D., Raymond J. and Bloch W. (1992) Prevention of pre-PCR mis-priming and primer dimerization improves low-copy-number amplifications. *Nucl. Acids Res.* **20**: 1717) and PCR was performed for 25 cycles on a TRIO-Thermoblock Thermal Cycler (Biometra, Göttingen, Germany). Each cycle consisting of a denaturation step at 94°C for 1 min, a primer annealing step at 70°C for 1 min, and an extension step at 72°C for 2 mins. This was followed by a single cycle consisting of a further elongation step at 72°C for 10 mins followed by cooling at 4°C. The PCR products were then extracted and purified from a TAE 1.4% standard agarose gel using a QIAquick™ gel extraction kit, following the protocol supplied by the manufacturer

(QIAGEN Ltd., UK). The PCR product V_HHe was then restriction enzyme digested with MscI and HindIII and ligated into reshF19Hc cloned in pg1d105 previously digested with the same two restriction enzymes. The MscI restriction recognition site is unique to all the reshaped human F19 V_H region versions and is not present in the pg1d105 expression vector. The HindIII restriction recognition site is a unique site in pg1d105 for clotting of V_H immunoglobulin genes.

[0093] Electroporation-competent XL-1 Blue E. coli cells were transformed with 1 µl of the ligated DNA and plated on agarose plates containing Ampicillin. Colonies were then screened for the presence and correct size of inserts by direct PCR on colonies (Güssow D. and Clackson T. (1989) Direct clone characterization from plaques and colonies by the polymerase chain reaction. *Nucl. Acids Res.* **17**: 4000) with primers HCMi and Hucγ1 hybridising to the flanking pg1d105 vector sequences (Table 5). DNA from positive colonies was prepared using a Plasmid Midi kit, following the protocol supplied by the manufacturer (QIAGEN Ltd., UK). DNA sequencing was performed by the dideoxy chain termination method (Sanger F., Nicklen S. and Coulson A. (1977) DNA sequencing with chain-terminating inhibitors. *Proc. natn. Acad. Sci. U. S. A.* **74**: 5463) directly from circular vector DNA using conventional heat denaturation (Andersen A., Pettersson A. and Kieldsen T. (1992) A fast and simple technique for sequencing plasmid DNA with sequenase using heat denaturation. *Biotechniques* **13**: 678) and Sequenase 2.0 (USB, Cleveland, OH). The DNA sequences of reshF19He is shown in Figure 32.

TABLE 4: PCR primers for the construction of reshaped human F19 heavy chain variable regions versions a to d.

1. Primers for the synthesis of version "a"

F19vha1 (47mer):

5' GTGTATTCAGTGAAGGTGTATCTACTAGTTTTACAGCTGACTTTCAC 3'

F19vha2 (53 mer):

5' TAGTAGATACACCTTCACTGAATACACCATACACTGGGTTAGACAGG

CCCCTG 3'

F19vha3 (71 mer):

5' CCCTTGAACCTCTGGTTGTAGTTAGGAATACCATTGTTAGGATTAATACC
TCCTATCCACTCCAGCCTTTG 3'

F19vha4 (71 mer):

5' TAACTACAACCAGAAGTTCAAGGGCCGGGCCACCTTGACCGTAGGCAA
GTCTGCCAGCACCGCCTACATGG 3'

F19vha5 (63 mer):

5' GCATGGCCCTCGTCGTAACCATAGGCGATTCTTCTTCTGGCGCAGTAGT
AGACTGCAGTGTCC 3'

F19vha6 (48 mer):

5' CTATGGTTACGACGAGGGCCATGCTATGGACTACTGGGGTCAAGGAAC 3'

2. Primers for the synthesis of version "c"

F19vhc4 (71 mer):

5' TAACTACAACCAGAAGTTCAAGGGCCGGGTCACCATCACCGTAGACA
CCTCTGCCAGCACCGCCTACATGG 3'

3. Primers for the synthesis of version "b" and "d"

F19vhbd6 (27 mer):

5' GGACACTGCAGTCTACTTCTGCGCCAG 3'

4. Primers hybridizing to the flanking PUC19 vector sequences

APCR1 (17 mer, sense primer): 5' TACGCAAACCGCCTCTC 3'

APCR4 (18 mer, anti-sense primer): 5' GAGTGCACCATATGCGGT 3'

TABLE 5: PCR primer for the construction of reshaped human F19 heavy chain variable regions version e

1. Primer for the synthesis of version "e"

F19MscIHe (65 mer, anti-sense):

5' CCTT TGGCCAGGGGCCTGTCTAACCCAGTGTATGGTGTATTCAGTGAAGGTG
MscI
 TATCCACTAGTTTCCACTAGTTT 3'

2. Primers hybridizing to the flanking pg1d105 mammalian expression vector sequences

HCMi (28 mer, sense): 5' GTCACCGTCCTTGACACGCGTCTCGGGA 3'

Hucy1 (17 mer, anti-sense): 5' TTGGAGGAGGGTGCCAG 3'

Example 5: Reshaped human F19 antibody concentrations in COS cells supernatants

[0094] COS cells were transfected with one pair of a series of reshaped human F19 antibody constructs and the human antibody concentration was measured using the IgG1/Kappa ELISA as described in example 2.

TABLE 6

Reshaped human F19 antibody concentrations in COS cell supernatants		
Transfected Antibody components		Human γ1/K
Heavy chain	Kappa light chain	concentration [μg/ml]
Ha	La	2.50
Ha	Lb	0.18
Hb	La	1.25
Hb	Lb	0.10
Hd	La	1.15
Hd	Lb	0.18
Ha	La	1.50
Ha	Lc	1.56

TABLE 6 (continued)

Reshaped human F19 antibody concentrations in COS cell supernatants		
Transfected Antibody components		Human $\gamma 1/K$
Heavy chain	Kappa light chain	concentration [$\mu\text{g/ml}$]
Hc	La	1.47
Hc	Lc	1.97
cF19	La	1.54
cF19	Lb	0.07
cF19	Lc	2.14

TABLE 7

Reshaped human F19 antibody concentrations in COS cell supernatants		
Transfected Antibody components		Human $\gamma 1/K$
Heavy chain	Kappa light chain	concentration [$\mu\text{g/ml}$]
Ha	La	2.00
Ha	Lc	2.50
Hc	La	2.90
Hc	Lc	3.00
He	La	2.80
He	Lc	3.50

RNA splicing events required for the expression of immunoglobulin genes in mammalian cells

[0095] Both mammalian expression vectors pKN100 and pg1d105 have an intron between the variable and the constant regions which is removed during the process of gene expression to give rise to an messenger RNA. The splicing event which consists of a DNA recombination between the heavy or light chain splice donor sites and the immunoglobulin splice acceptor site is described in Figure 34.

Example 6: Flow cytometric analysis of the binding of cF19 and $L_A H_C$ to FAP-expressing human cells

[0096] The ability of $L_A H_C$ to bind to both recombinant and endogenously expressed FAP on cell surface was tested.

[0097] The example was conducted to determine the binding of $L_A H_C$ to cellular FAP. Both naturally FAP expressing MF-SH human tumour cells and FAP-transfected human tumour cell lines were used as cellular targets. $L_A H_C$ was studied in cytofluorometric assays evaluating direct binding to target cells as well as by the inhibitory effect on the binding of either murine F19 or chimeric cF19 anti-FAP antibodies.

[0098] Antibodies and cell lines used were F19 (murine monoclonal anti-human FAP antibody, IgG1 subclass), mIgG (murine immunoglobulin, IgG class), cF19 (chimeric monoclonal anti-human FAP antibody, IgG1 subclass), $L_A H_C$ (reshaped monoclonal anti-human FAP antibody, IgG1 subclass), hIgG1 (human immunoglobulin, IgG1 subclass), MF-SH (human malignant fibrous histiocytoma cell line), HT-1080 (human fibrosarcoma cell line), HT-1080FAP clone 33 (HT-1080 cell line transfected with cDNA encoding human FAP)

Direct binding of L_AH_C to FAP on the surface of human tumour cell lines

[0099] 5x10⁵ cells of the tumour cell line under investigation were incubated with the indicated concentration of test or control antibody in a total volume of 0.2 ml phosphate-buffered saline (PBS) supplemented with 1% bovine serum albumin (BSA) for 30 min on ice.

[0100] Subsequently, cells were washed twice with 2 ml of PBS, resuspended in 0.2 ml of PBS supplemented with 1% BSA, the appropriate anti-Ig-antibody as secondary reagent (either a 1:20 dilution of goat anti-mouse Ig FITC-labeled [Dianova] or a 1:20 dilution of mouse anti-human IgG FITC-labeled [Dianova]) and incubated for another 30 min on ice.

[0101] Cells were again washed twice with 2 ml of PBS, resuspended in a total volume of 0.5 ml of PBS supplemented with 1% paraformaldehyde (PFA) and kept on ice. Single cell fluorescence was determined cytofluorometrically by analysing the cellular green fluorescence in the 488nm light of an EPICS XL (Coulter).

Inhibitory effect of L_AH_C on binding of biotinylated cF19 to FAP on the surface of human cell lines

[0102] 5x10⁵ cells of the tumour cell line under investigation were incubated with the indicated concentration of the biotin-labelled antibody in a total volume of 0.2 ml PBS supplemented with 1% BSA and the simultaneously added unlabelled test or control antibody for 30 min on ice. Subsequently, cells were washed twice with 2 ml of PBS, resuspended in 0.2 ml of PBS supplemented with 1% BSA, 1:40 diluted streptavidin-FITC (Dianova) as secondary reagent and incubated for another 30 min on ice.

[0103] Alternatively, cells were incubated with the indicated concentrations of murine F19 and cell-bound antibody detected via 1:20 diluted goat anti-mouse Ig labelled with FITC by comparable incubation steps.

[0104] In each case, cells were finally washed twice with 2 ml of PBS, resuspended in a total volume of 0.5 ml PBS supplemented with 1% PFA and kept on ice. Single cell fluorescence was determined cytofluorometrically by analysing the cellular green fluorescence in the 488nm light of an EPICS XL (Coulter).

[0105] Both, cF19 and L_AH_C bind in a concentration dependent manner specifically to FAP-transfected HT-1080FAP clone33 human tumour cells (Table 8). No binding to FAP-negative HT-1080 cells was detectable (Table 9). Both cF19 and L_AH_C bound in a concentration dependent manner to human MF-SH cells endogenously expressing FAP (Table 10).

[0106] Biotinylated cF19 in a concentration dependent manner bound to human HT-1080FAP clone 33 (Table 11). No binding was detectable to FAP-negative HT-1080 cells (Table 12).

[0107] Binding of biotinylated cF19 to HT-1080FAP clone 33 cells was inhibited by both unlabelled cF19 and unlabelled L_AH_C (Table 13).

[0108] Chimeric anti-human FAP monoclonal antibody cF19 as well as reshaped human anti-human FAP monoclonal antibody L_AH_C (example 10) were shown to bind directly to FAP expressed on human cell lines either endogenously expressing this protein or transfected with cDNA encoding for it. This binding was shown to be concentration dependent. Binding of biotinylated cF19 could be inhibited by both unlabelled cF19 and unlabelled L_AH_C.

[0109] Using cytofluorometric technology, direct binding as well as inhibition of specifically binding reagents showed specificity of chimeric cF19 and reshaped L_AH_C human monoclonal antibodies to cell surface expressed FAP.

Table 8

Binding of anti-FAP antibodies to HT-1080FAP clone 33 cells			
Concentration of anti-body	Mean fluorescence intensity		
[ng/mL]	hIgG1	cF19	L _A H _C
500.0	0.12	6.65	2.76
100.0	0.12	1.63	0.66
20.0	0.12	0.43	0.22
4.0	0.12	0.17	0.15
0.8	0.12	0.14	0.13

Table 9

Binding of anti-FAP antibodies to non-transfected HT-1080 cells			
Concentration of anti-body	Mean fluorescence intensity		
[ng/mL]	hlgG1	cF19	L _A H _C
500.0	0.11	0.11	0.12
100.0	0.11	0.11	0.11
20.0	0.11	0.11	0.12
4.0	0.11	0.11	0.12
0.8	0.11	0.11	0.11

Table 10

Binding of anti-FAP antibodies to MF-SH cells			
Concentration of anti-body	Mean fluorescence intensity		
[ng/mL]	hlgG1	cF19	L _A H _C
4.0	0.6	3.6	2.8
2.0	n.d.	3.3	2.5
1.0	n.d.	2.4	1.9
0.5	n.d.	1.8	1.3

n.d.: not done

Table 11

Binding of biotinylated cF19 antibody to HT-1080FAP clone 33 cells		
Concentration of anti-body	Mean fluorescence intensity	
[ng/ml]	Biotinylated hlgG1	Biotinylated cF19
5,000.0	0.2	36.5
1,000.0	0.2	18.1
200.0	0.2	4.5
40.0	0.2	1.3
8.0	0.2	0.5
1.6	0.3	0.3

Table 12

Binding of biotinylated cF19 antibody to non-transfected HT-1080 cells		
Concentration of anti-body	Mean fluorescence intensity	
[ng/ml]	Biotinylated hIgG1	Biotinylated cF19
5,000.0	0.1	0.1
1,000.0	0.1	0.1
200.0	0.1	0.1
40.0	0.1	0.1
8.0	0.1	0.1
1.6	0.1	0.1

Table 13

Competition of anti-FAP antibodies with the binding of biotinylated cF19 to HT-1080FAP clone 33 cells		
	Concentration of competitor antibody	Mean fluorescence concentration
Competitor antibody	[$\mu\text{g}/\text{mL}$]	
no	0.00	11.2
hIgG1	1.00	9.0
hIgG1	3.16	11.3
hIgG1	10.00	9.8
hIgG1	31.66	10.3
cF19	1.00	7.5
cF19	3.16	4.8
cF19	10.00	1.3
cF19	31.66	1.2
L _A H _C	1.00	8.0
L _A H _C	3.16	5.5
L _A H _C	10.00	2.9
L _A H _C	31.66	1.7
Biotinylated cF19 was used at a concentration of 1 $\mu\text{g}/\text{mL}$ in all tests shown in the table.		

Example 7: In vitro immune effector functions of monoclonal antibody L_AH_C

[0110] This experiment was conducted to determine the potential of the monoclonal antibody (mab) L_AH_C with specificity for fibroblast activation antigen (FAP) to lyse FAP-expressing targets in the presence of human complement or human mononuclear leukocytes, respectively.

[0111] In particular, the ability of L_AH_C to mediate cytotoxic effects against HT-1080FAP clone 33 cells, which expressed human FAP on the surface, was studied. Cytotoxicity was determined in vitro using the following approach: ^{51}Cr -labelled target cells were incubated in the presence of L_AH_C with human serum as source of complement or human MNC (peripheral blood mononuclear cells) as effectors. Release of ^{51}Cr was measured as measure of target-cell lysis.

[0112] Antibodies and cell lines used were L_AH_C (reshaped human anti-human FAP IgG1 antibody), hIgG1 (human IgG1 isotype control), 3S193 (murine monoclonal anti-Lewis Y IgG3 antibody), mIgG (murine IgG control), HT-1080 (human fibrosarcoma), HT-1080FAP clone 33, (HT1080 transfected with cDNA encoding human FAP), MCF-7 (human breast adenocarcinoma cell line).

Complement-mediated lysis of target cells by L_AH_C

[0113] Tumour cells were radiolabelled by incubation in RPMI1640 medium with 100 μ l ^{51}Cr (NEN) at 37° C for one hour. Subsequently, cells were washed twice in ^{51}Cr -free medium and resuspended at a concentration of 2×10^5 cells per mL.

[0114] Human serum as source of complement was freshly prepared from blood of different volunteers. Blood was taken by puncturing the arm vein, remained at room temperature for one hour to allow clotting to occur, and was kept at 4° C over night. Serum was separated by centrifugation and taken off from the sediment.

[0115] The antibody under study was diluted from the stock solution to the appropriate concentration in RPMI1640 cell culture medium.

[0116] 1×10^4 radiolabelled tumour cells of the indicated cell line were incubated in the presence of different concentrations of test or control antibody and 25% of the human serum used as source of complement for 2 h at 37° C in a 95% air and 5% CO $_2$ incubator. Incubation was performed in U-shaped 96-well plates in a total volume of 200 μ l RPMI1640 and done in triplicate. After the incubation period, plates were centrifugated, 100 μ l of the supernatant were taken off and radioactivity was determined in a gamma-counter. Total number of incorporated radioactivity was determined by measuring 10^4 target cells. Spontaneous release was defined as activity released from the target cells in the absence of both antibody and complement during the described incubation period.

[0117] Specific lysis was calculated as follows:

$$\text{specific lysis (in \%)} = \frac{[\text{activity sample}] - [\text{activity spontaneous release}]}{[\text{maximum activity}] - [\text{activity spontaneous release}]} \times 100$$

Antibody-dependent cellular cytotoxicity (ADCC) of L_AH_C

[0118] Tumour cells were radiolabelled by incubation in RPMI1640 medium with 100 μ l ^{51}Cr at 37°C for one hour. Subsequently, cells were washed twice in ^{51}Cr -free medium and resuspended at a concentration of 2×10^5 cells per mL.

[0119] MNC (peripheral blood mononuclear cells) were prepared from peripheral blood taken by puncturing the arm vein of different healthy human volunteers. Clotting was prevented by the addition of 20% citrate buffer. MNC from 4 mL of this blood preparation were purified by centrifugation (30 min at 400 G and room temperature) on 3 mL of lymphocyte preparation medium (Boehringer Mannheim, Germany). MNC (peripheral blood mononuclear cells) were taken off from the gradient, washed three times and diluted with RPMI1640 to the appropriate concentration. Lymphocyte activated killer (LAK) cells were derived from MNC (peripheral blood mononuclear cells) by incubation for 5 days at 37° C in a 95% air and 5% CO $_2$ incubator at an initial density of 1.3×10^6 cells per mL in the presence of 100U recombinant human Interleukin-2 (IL-2). The antibody under study was diluted from the stock solution to the appropriate concentration in RPMI1640 cell culture medium.

[0120] 1×10^4 radiolabelled tumour cells of the indicated cell line were incubated for 5 h at 37°C and 5%CO $_2$ in the presence of different concentrations of test or control antibody and MNC (peripheral blood mononuclear cells) in a number necessary to reach the indicated effector:target cell ratio. Incubation was performed in U-shaped 96-well plates in a total volume of 200 μ l RPMI1640 and done in duplicate.

[0121] After the incubation period, plates were centrifugated, 100 μ l of the supernatant were taken off and radioactivity was determined in a gamma-counter. Total number of incorporated radioactivity was determined by measuring 10^4

target cells. Spontaneous release was defined as activity released from the target cells in the absence of both antibody and effector cells during the described incubation period.

[0122] Specific lysis was calculated as follows:

$$\text{specific lysis (in \%)} = \frac{[\text{activity sample}] - [\text{activity spontaneous release}]}{[\text{maximum activity}] - [\text{activity spontaneous release}]} \times 100$$

Antibody mediated complement lysis of tumour cells

[0123] No complement mediated lysis above control was seen in HT-1080FAP clone 33 cells with L_AH_C up to a concentration of 50 $\mu\text{g/mL}$ (Table 14, Table 15a)

[0124] Lytic activity of human serum used as source of complement was shown by lysis of MCF-7 human breast carcinoma cells in the presence of 12.5 $\mu\text{g/mL}$ 3S193, a murine monoclonal anti-Lewis^y antibody with known complement activating ability (Table 15b)

Antibody mediated cellular lysis of tumour cells

[0125] In the presence of L_AH_C in a concentration of up to 10 $\mu\text{g/mL}$, no lysis of HT-1080FAP clone 33 above isotype control was detectable in ADCC mediated by human MNC (peripheral blood mononuclear cells, Table 16) or human LAK cells (lymphokine activated killer cell) (Table 17) at an effector:target ratio of 50:1:

[0126] In appropriate in vitro assays with either human complement or with human MNC (peripheral blood mononuclear cells) as effector mechanisms, human anti-FAP monoclonal antibody L_AH_C revealed no relevant cytotoxic effect above controls on FAP expressing tumor cell line HT-1080FAP clone 33.

[0127] In vitro, L_AH_C is unable to mediate cytotoxicity effected by human complement or human MNC (peripheral blood mononuclear cells) on a cell line positive for FAP, the antigen recognized by this antibody.

Table 14

Specific complement lysis (in %) of HT-1080FAP clone 33 tumor cell targets mediated by L_AH_C		
Source of human serum:	HT-1080 clone 33:	
concentration of anti-body	hIgG1 isotype control	L_AH_C
A 50 $\mu\text{g/mL}$	5	4
A 10 $\mu\text{g/mL}$	5	3
B 50 $\mu\text{g/mL}$	7	5
B 10 $\mu\text{g/mL}$	6	5
0 $\mu\text{g/mL}$	0	0
Incubation: 2 hours at 37°C, 25% serum from human volunteers A or B, respectively, as source of complement.		

Table 15a

Specific complement lysis (in %) of HT-1080FAP clone 33 tumor cell targets mediated by human anti-FAP monoclonal antibody L_AH_C		
Source of human serum:	HT1080clone 33:	
concentration of anti-body	hIgG1	L _A H _C
A 10.00 µg/ml	2	1
A 2.50 µg/ml	2	2
A 0.60 µg/ml	1	1
A 0.15 µg/ml	1	2
A 0.00 µg/ml	2	2
B 10.00 µg/ml	2	2
B 2.50 µg/ml	2	2
B 0.60 µg/ml	2	2
B 0.15 µg/ml	2	2
B 0.00 µg/ml	2	2
C 10.00 µg/ml	2	2
C 2.50 µg/ml	1	1
C 0.60 µg/ml	1	1
C 0.15 µg/ml	2	1
C 0.00 µg/ml	3	3
Incubation: 2 hours at 37°C, 25% serum from human volunteers A, B or C, respectively, as source of complement.		

Table 15b

Specific complement lysis (in %) of MCF-7 tumour cell targets mediated by murine anti-Lewis^y monoclonal antibody 3S193		
Source of human serum:	MCF-7:	
concentration of anti-body	mIgG	3S193
A 10.00 µg/ml	0	21
A 2.50 µg/ml	1	21
A 0.60 µg/ml	0	21
A 0.15 µg/ml	1	18
A 0.00 µg/ml	0	0
B 10.00 µg/ml	1	13
B 2.50 µg/ml	0	17

Table 15b (continued)

Specific complement lysis (in %) of MCF-7 tumour cell targets mediated by murine anti-Lewis^y monoclonal antibody 3S193		
Source of human serum:	MCF-7:	
concentration of anti-body	mlgG	3S193
B 0.60 µg/ml	1	18
B 0.15 µg/ml	1	15
B 0.00 µg/ml	0	0
C 10.00 µg/ml	1	22
C 2.50 µg/ml	0	23
C 0.60 µg/ml	1	26
C 0.15 µg/ml	1	20
C 0.00 µg/ml	1	1
Incubation: 2 hours at 37° C, 25% serum from human volunteers A, B or C, as source of complement.		

Table 16

ADCC (antibody-dependant cellular cytotoxicity) (specific lysis in %) of HT-1080FAP clone 33 target cells by human MNC (peripheral blood mononuclear cells) mediated by L_AH_C.		
HT-1080FAP clone 33:		
Concentration of anti-body:	HT-1080FAP clone 33:	
[in µg/mL]	hlgG1	L _A H _C
10.000	2	2
2.500	2	2
0.625	2	2
0.156	3	3
0.000	3	3
Incubation: 5 hours at 37°C, 10 ⁴ target cells and an effector:target cell ration of 50:1.		

Table 17

ADCC (antibody-dependant cellular cytotoxicity, specific lysis in %) of HT-1080FAP clone 33 target cells by LAK cells (lymphokine activated killer cells) mediated by L _A H _C .		
Concentration of anti-body:	HT-1080FAP clone 33:	
[in µg/mL]	hIgG1	L _A H _C
10.000	12	14
2.500	14	17
0.625	14	21
0.156	15	21
0.000	14	14
Incubation: 5 hours at 37°C, 10 ⁴ target cells and an effector:target cell ration of 50:1.		

Example 8: Immunohistochemical analysis of monoclonal antibody L_AH_C binding to normal and neoplastic human tissues

[0128] This experiment was performed to determine the binding characteristics of the humanized mAb L_AH_C to normal and neoplastic human tissues.

[0129] The following antibodies were used: L_AH_C, cF19, and the negative control hu IgG1 were directly biotinylated according to methods of the state of the art and used at concentrations of 2.5 to 0.25 mg/ ml in 2% BSA/PBS (bovine serum albumin in phosphate-buffered saline). Murine mAb F19 was used as tissue culture supernatant of the F19 hybridoma, at dilutions of 1:5 to 1:10 in 2% BSA/PBS.

[0130] The following reagents were used for immunochemical assays: Streptavidin peroxidase complex (Vector Labs., Burlingame, CA, USA), Avidin-biotin peroxidase complex (Vector Labs.), Biotinylated horse anti-mouse (Vector Labs.), DAB (diaminobenzidine, Sigma Chemical Co. St. Louis, MO, USA), Harris' hematoxylin.

[0131] Fresh frozen tissue samples examined included the following: Normal colon, breast, lung, stomach, pancreas, skin, larynx, urinary bladder, smooth and skeletal muscle.

[0132] Among the tumors tested were carcinomas from breast, colon, lung, esophagus, uterus, ovary, pancreas, stomach, and head and neck.

[0133] An indirect immunoperoxidase method was carried out according to state of the art methods (Garin-Chesa P, Old LJ, Rettig WJ: Cell surface glycoprotein of reactive stromal fibroblasts as a potential antibody target in human epithelial cancers. Proc Natl Acad Sci USA 1990; 87:7235-7239) on five micrometer thickness fresh frozen sections.

[0134] DAB was used as a substrate for the final reaction product. The sections were counterstained with Harris' hematoxylin and examined for antigen expression.

L_AH_C expression in normal human tissues

[0135] The normal tissues tested were negative for L_AH_C expression, except for the normal pancreas in which a subset of positive endocrine cells in the islets of Langerhans (A cells) were identified with L_AH_C, cF19 and F19. (Table 18). No immunoreactivity was observed with the hu IgG1 (human immunoglobulin IgG1 subclass) used as a negative control.

L_AH_C expression in tumors

[0136] In the tumor samples, L_AH_C, cF19 and F19 showed an indistinguishable pattern of expression in the tumor stromal fibroblasts. A strong and homogeneous expression was found in the majority of the cases examined, especially in the cancer samples derived from breast, colon, lung, pancreas and in the squamous cell carcinomas (SQCC) of the head and neck tested (Table 19). No immunoreactivity was observed with the hu IgG1 used as negative control.

[0137] L_AH_C, cF19 and F19 showed immunoreactivity with the tumor stromal fibroblasts in the epithelial cancer samples tested. No L_AH_C or F19 immuno-reactivity was seen with either the fibrocytes of the normal organ mesenchyme or

the parenchymal cells of normal adult organs. The only exception was a subset of endocrine cells in the pancreatic islets, presumably glucagon-producing A cells, which react with the anti-FAP antibodies.

[0138] Immunohistochemical analysis of L_AH_C in normal human tissues and FAP-expressing human carcinomas showed indistinguishable patterns of binding for L_AH_C , cF19 and murine mAb F19.

Table 18

Immunoreactivity of mAbs L_AH_C , cF19 and F19 with normal human tissues				
Tissue type		L_AH_C	cF19	F19
Breast	-Duct epithelium	-	-	-
	-Myoepithelial cells	-	-	-
Colon	-Glandular epithelium	-	-	-
	-Smooth muscle	-	-	-
Lung	-Bronchial epithelium	-	-	-
	-Alveolar epithelium	-	-	-
Stomach	-Glandular epithelium	-	-	-
	-Smooth muscle	-	-	-
Urinary bladder		-Urothelium	-	-
		-Smooth muscle	-	-
Pancreas	-Exocrine acini	-	-	-
	-Endocrine islet cells	+ subset only	+subset only	+ subset only
Larynx -Squamous epithelium		-	-	-
Lymph node -Lymphocytes		-	-	-
Skeletal muscle-		-	-	-
Connective tissue		-	-	-
Skin	-Keratinocytes	-	-	-
	-Sweat glands	-	-	-

Table 19

Immunoreactivity of mAbs L_AH_C , cF19 and F19 with human tumor samples				
Tumor type	No.	L_AH_C	cF19	F19
Breast cancers (infiltrating ductal type)	7	7 Positive	7 Positive	7 Positive
Colon cancers (adenocarcinomas)	7	7 Positive	7 Positive	7 Positive
Lung carcinomas (adenocarcinoma (2) large cell type (2) squamous type (4)	8	7 Positive	7 Positive	7 Positive
		1 Negative	1 Negative	1 Negative
Esophageal cancers (squamous type)	1	1 Positive	1 Positive	1 Positive
Endometrial cancers (adenocarcinoma)	1	1 Negative	1 Negative	1 Negative
Gastric cancers (adenocarcinoma)	2	2 Negative	2 Negative	2 Negative
Ovarian cancers (serous denocarcinoma)	2	1 Positive	1 Positive	1 Positive
		1 Negative	1 Negative	1 Negative

Table 19 (continued)

Immunoreactivity of mAbs L _A H _C , cF19 and F19 with human tumor samples				
Tumor type	No.	L _A H _C	cF19	F19
Pancreatic cancers (adenocarcinomas)	2	2 Positive	2 Positive	2 Positive
Head and neck cancers (squamous cell type)	4	4 Positive	4 Positive	4 Positive
Abbreviations: No, number of cases from different patients studied; positive, number of cases showing antigen expression in the tumor stroma; negative, number of cases tested that lacked detectable antigen expression.				

Example 9: Species specificity of L_AH_C binding in tissue sections

[0139] This experiment was conducted to assess the reactivity of L_AH_C with tissues from mouse, rat, rabbit and cynomolgus monkeys by immunohistochemical methods.

[0140] Also used in these tests were cF19 and hulgG1 as negative controls. The reagents used for immunohistochemistry were Streptavidin peroxidase complex (Vector Labs., Burlingame, CA, USA), DAB (Sigma Chemical Co., St. Louis, MO, USA) and Harris' hematoxylin.

[0141] The following fresh frozen tissue samples from mouse, rat, rabbit and cynomolgus were tested: Brain, liver, lung, kidney, stomach, pancreas, intestine, thymus, skin, muscle, heart, spleen, ovary, uterus and testes. As positive control, sections from normal human pancreas and a breast carcinoma sample were included in every assay.

Immunohistochemistry

[0142] An indirect immunoperoxidase method was carried out as described in the state of the art (Garin-Chesa P, Old LJ, Rettig WJ: Cell surface glycoprotein of reactive stromal fibroblasts as a potential antibody target in human epithelial cancers. Proc Natl Acad Sci USA 1990; 87:7235-7239) on five micrometer thickness fresh frozen sections. The antibodies L_AH_C, cF19 and hu IgG1 (at 1 µg/ml) were biotinylated according to the state of the art and were detected with streptavidin peroxidase complex. DAB was used as a substrate for the final reaction product. The sections were counterstained with Harris' hematoxylin and examined for antigen expression.

[0143] The normal tissues tested did not react with either L_AH_C or cF19 in the experiments (Table 1).

[0144] The normal human pancreas used as positive control showed L_AH_C and cF19 binding in a subset of endocrine cells in the islets of Langerhans as previously described for F19. In addition, binding of L_AH_C and cF19 was seen in the tumor stromal fibroblasts in the breast carcinoma sample.

[0145] Immunohistochemical analysis of normal tissues from mouse, rat, rabbit and cynomolgus failed to detect any binding of either L_AH_C or cF19, in the experiments performed.

Table 20

Binding of L _A H _C to tissue sections of non-human species, as determined by immunohistochemistry.					
Organ / Tissue typ		Mouse	Rat	Rabbit	Cynomolgus
Brain	-Cerebral cortex	-	-	-	
	-Cerebellum	-	-	-	-
Liver	-Hepatocytes	-	-	-	-
	-Portal triad	-	-	-	-
Lung	-Bronchi	-	-	-	-
	-Alveoli	-	-	-	-
Kidney	-Glomeruli	-	-	-	-
	-Tubular epithelium	-	-	-	-
Stomach	-Glandular epithelium				
	-Smooth muscle	-	-	-	-
Pancreas	-Exocrine acini	-	-	-	-
	-Endocrine islets	-	-	-	-
Intestine	-Glandular epithelium	-	-	-	-
	-Smooth muscle	-	-	-	-
Thymus -Lymphocytes		-	-	-	-
Skin	-Keratinocytes	-	-	-	-
	-Sweat glands	-	-	-	-
	-Hair follicles	-	-	-	-
Skeletal muscle		-	-	-	-
Heart		-	-	-	-
Spleen -Lymphocytes		-	-	-	-
Ovary	-Follicular epithelium	-	-	-	-
	-Stroma	-	-	-	-
Uterus	-Myometrium	-	-	-	-
	-Cervix uteri	-	-	-	-
Testis -Tubular epithelium		nt	nt	nt	-
Connective tissue		-	-	-	-

nt, not tested

Example 10: Construction of cell lines producing chimeric and reshaped anti-FAP monoclonal antibodies

[0146] The objective of this experiment was to demonstrate stable cell lines according to the invention expressing L_AH_C, L_AH_A, L_BH_B, L_BH_D, and cF19 in CHO DG44 cells. Stable cell lines transfected with humanized or chimeric F19 antibodies were produced and their identity was confirmed by PCR amplification of heavy and light variable regions using genomic DANN derived from each transfectant as template.

[0147] CHO DG44 cells maintained under serum-free conditions in SFM-II medium. Lipofectin and SFM-II serum-free medium were obtained from Gibco/BRL. Geneticin and all restriction enzymes were obtained from Boehringer Mannheim. Pfu polymerase was obtained from Stratagene.

[0148] DNA for transfections was purified from E. coli cells using QiaFilter Maxi Cartridges (Qiagen) as directed by the manufacturer. All DNA preparations were examined by restriction enzyme digestion. Sequences of L_AH_C variable regions in their respective vectors were confirmed using an ABI PRISM 310 Sequencer.

[0149] Further information regarding the vectors and DNA sequences employed is available in the prior examples.

Transfection of CHO DG44 cells

[0150] Cells in logarithmic growth were plated into 6 well plates containing 1 mL fresh SFM-II medium. Plasmids encoding heavy and light chains of humanized or chimeric F19 versions were cotransfected into CHO DG44 cells using liposomal transfection. Liposomes were prepared using 6 μ L Lipofectin reagent and 0.5 μ g of each vector (one for the desired heavy chain and one for the light) as described for LipofectAMINE transfections except that SFM-II medium was used to dilute all reagents. Twenty-four hours later, cells were diluted 1:10 into SFM-II medium containing 300 μ g/mL Geneticin. After the initial phase of cell killing was over (10-14 days), the concentration of Geneticin was reduced to 200 mg/mL and methotrexate was added to a final concentration of 5 nM. Methotrexate concentrations were increased after 10-14 days to a final concentration of 20 nM.

PCR Amplification of transfectant DNA

[0151] 10^7 CHO DG44 cells were centrifuged in an Eppendorf microcentrifuge briefly at full speed, washed once with PBS, and pelleted once again. Genomic DNA was prepared by ethanol precipitation after SDS lysis and Proteinase K treatment of the cell pellets.

[0152] A mixture containing one of the following primer pairs, dNTPs, buffer, and Pfu polymerase was used to amplify either the heavy or light chain variable region using genomic DNA as template. The resulting PCR products were digested with the appropriate restriction enzyme and analyzed by agarose gel electrophoresis to confirm their identity.

Light chain primer set:

[0153]

5'-GAG ACA TTG TGA CCC AAT CTC C - 3' PKN 1690

5'- GAC AGT CAT AAA CTG CCA CAT CTT C - 3' PKN.1930.R

Heavy chain primer set:

[0154]

5'-TTG ACA CGC GTC TCG GGA AGC TT - 3' PG 5863

5'- GGC GCA GAG GAT CCA CTC ACC T - 3' PG 6332.R

[0155] The undigested heavy chain PCR product has a predicted size of 469 bp while the light chain PCR product has a predicted size of 286 bp. Verification of identity was determined by restriction enzyme digest with BstEII (heavy chain) or NlaIV (light chain).

[0156] CHO cell lines were transfected with L_AH_C , L_AH_A , L_BH_B , L_BH_D , as well as cF19. Geneticin-resistant cells were obtained and these cells were further selected for resistance to methotrexate. PCR amplification of the light and heavy chain DNA produced the expected bands and confirmed the identity of L_AH_C , L_AH_A and L_BH_D transfectants. The L_AH_C full length heavy chain PCR product was subcloned and resequenced in its entirety.

[0157] The cells described were maintained under serum-free conditions at all times and were not treated with animal-derived products such as trypsin.

[0158] Producer cell lines transfected with expressing monoclonal L_AH_C , L_AH_A , L_BH_B , L_BH_D and cF19 antibodies were produced. Their identities were confirmed using PCR amplification of both their heavy and light chain variable regions. The DNA sequence of the heavy chain variable region PCR products for L_AH_C -transfected cells was confirmed.

Example 11: Expression of antibody proteins in Chinese hamster ovary DG 44 cells and their purification

[0159] The objective of this experiment was to express and purify of L_AH_C , L_AH_A , L_BH_B , and L_BH_D mAbs to enable their characterization. Other goals included the establishment of a quantitative ELISA to permit measurement of anti-

body concentrations in both crude media samples as well as purified Ig samples and determination of relative expression levels of various humanized F19 constructs using this assay.

[0160] Serum-free CHO DG44 cells and USP-grade methotrexate were obtained from the Biotechnical Production Unit of the Dr. Karl Thomae GmbH, Biberach, Germany; both products are also commercially available. Cells were maintained under serum-free conditions at all times. SFM-II serum-free medium was obtained from Gibco/BRL.

[0161] Protein A agarose was from Pierce Chemical (Indianapolis, IN, USA). Human IgG1 standards (Cat. No. I 3889), p-Nitrophenyl phosphate tablets (N 2640), bovine serum albumin (BSA) (A 7906), and goat anti-human kappa chain specific alkaline phosphatase-conjugated antibody (A 3813) were obtained from Sigma Chemical (St. Louis, MO, USA). Goat anti-human gamma-chain specific alkaline phosphatase-conjugated antibody was obtained from Jackson ImmunoResearch Laboratories (through Stratech Scientific). Tris-buffered saline (TBS) consisted of 150 mM NaCl, 50 mM Tris, pH 7.5.

Cell culture conditions for antibody expression

[0162] Cells were cultured and L_AH_C -producing cells were maintained in T-175 flasks in SFM-II serum-free medium without agitation. The medium contained 200 μ g/mL Geneticin and 20 nM methotrexate without antibiotics. Cells were passaged by dilution, were not adherent, and grew in small clusters. When the cells reached stationary phase, the medium was collected and centrifuged to remove cells and frozen at -20°C until needed.

Purification of L_AH_C

[0163] All purification steps were carried out at 4° C. A C10/10 column (Pharmacia Fine Chemicals) was packed with Protein A agarose (3 mL bed volume). The column was washed with TBS and preeluted once with 0.1 M Na citrate, pH 3.0 to insure that no loosely bound material remained on the column. The column was then immediately reequilibrated with TBS and stored at 4°C. Spent culture supernatants were thawed and centrifuged at 10,000 xg for 30 minutes prior to Protein A chromatography to remove debris and diluted with an equal volume of TBS. This material was loaded onto the Protein A column at 0.5 mL/min using a P-1 peristaltic pump (Pharmacia) and washed with TBS until the absorbance at 280 nm was undetectable. Elution of the antibody was initiated with 0.1 M Na citrate pH 3.0 at approximately 0.2 mL/min. The elution was monitored at 280 nm and one mL fractions of the eluted material were collected into tubes containing sufficient Tris base pH 9 to neutralize the citrate buffer. Protein-containing fractions were pooled and concentrated using an Amicon filtration apparatus with a YM-30 filter and dialyzed against PBS. The column was immediately regenerated with TBS. Protein dye-binding assays were performed with the BioRad (Hercules, California) protein determination kit, according to the manufacturer's instructions, using bovine serum albumin as a standard.

Human IgG (gamma immunoglobulin) ELISA

[0164] ELISA plates were coated overnight with 100 μ L of goat anti-human gamma-chain specific alkaline phosphatase-conjugated antibody at 0.4 mg/mL in coating buffer at 4°C. Coating antibody was removed and plates were blocked with 2% BSA in PBS for 2 hours. All subsequent steps were performed at 37°C. Blocking buffer was replaced with antibody samples or human IgG1 standard diluted in dilution buffer, serially diluted in a 200mL volume, and incubated for one hour. Negative controls included dilution buffer and/or culture medium of nontransfected cells. Wells were washed and 100 μ L of goat anti-human kappa chain specific alkaline phosphatase-conjugated antibody diluted 1:5000 was added and incubated for one hour. Wells were washed and 100 μ L reaction buffer was added and incubated for 30 minutes. The reaction was stopped by addition of 1 M NaOH and absorbance read at 405 nm in an ELISA plate reader. Results were analyzed by four-parameter iterative curve fitting.

[0165] Amino acid analysis was performed according to methods available in the state of the art.

[0166] Monoclonal antibody L_AH_C was produced and purified to homogeneity using Protein A affinity chromatography. ELISA assays using human IgG1 as standard indicated L_AH_C recoveries exceeding 70%. The purity of the material was estimated to be >90% by SDS-polyacrylamide gel electrophoresis. Representative expression data and typical purification yields are shown in Table 21.

Table 21

Expression data and purification yields FAP antibody proteins in CHO cells			
Antibody	Expression levels in crude media samples (ELISA)	Purified antibody yields	Yield improvement [purified antibody]
H _C L _A	7 - 10 mg/L	~ 5 - 7 mg/L	500 - 700
H _A L _A	5 - 7 mg/mL	~ 3 - 4 mg/L	300 - 400
H _B L _B	0.5 - 1 mg/mL	~ 0.2 - 0.5 mg/L	20 - 50
H _D L _B	0.8 - 1.5 mg/mL	~ 0.3 - 0.8 mg/L	30 - 60
Chimeric F19	~ 0.02 mg/mL	< 0.01 mg/L	1
Representative expression data for each of the anti-FAP antibodies produced in this study are shown. Recoveries after Protein A agarose affinity chromatography were based on protein dye-binding measurements of the purified Ig using BSA as a standard.			

Example 12: Binding of monoclonal antibody L_AH_C to isolated recombinant human FAP

[0167] The objective of this study was to characterize binding of L_AH_C to isolated recombinant human FAP.

CD8-FAP ELISA

[0168] ELISA plates were coated overnight with 100 µL of mouse anti-rat antibody (Sigma Chemical R0761) at 1:2000 in coating buffer at 4 °C. Coating antibody was removed and plates were blocked with 2% BSA in PBS for one hour. All subsequent steps were performed at room temperature. Blocking buffer was replaced with 100 mL of 1 µg/mL rat anti-CD8 antibody (Pharmingen 01041D) and incubated for one hour. Plates were washed and 100 µL CD8-FAP culture supernatant (1:2 in PBS) was added and allowed to bind for one hour. Plates were washed and antibody samples were added (two-fold serial dilutions) in a 100 µL volume and incubated for one hour. Negative controls included human IgG and/or culture medium of nontransfected cells. Wells were washed and 100 µL of horse radish peroxidase (HRP) conjugated mouse anti-human IgG1 antibody (Zymed 05-3320) diluted 1:500 in dilution buffer were added and incubated for one hour. Wells were washed and 100 µL HRP substrate, (azino-bis (3-ethylbenzthiazoline 6-sulfonic) acid, Sigma Chemical A9941), were added and incubated for 60 minutes. The reaction was stopped by addition of 1 M NaOH and absorbance read at 405/490 nm in an ELISA plate reader. Results were analyzed by four parameter curve iterative curve fitting.

[0169] Alternatively, plates were coated directly with cF19. FAP (recombinant human FAP) was allowed to bind to these plates as above and biotinylated L_AH_C (~1 µg/mL) was then added. Antibody binding was detected with HRP-streptavidin conjugate as above.

Solubilization of membrane-bound human FAP

[0170] FAP-expressing 293FAP 1/2 cells or control 293 cells were washed with PBS and lysed with 1% Triton X-114 in Tris-buffered saline. Nuclei and debris were removed by centrifugation at 10,000 xg. The supernatant was phase-partitioned (Estreicher A, Wohlend A, Belin D, Scheuning WD, Vasalli JD. Characterization of the cellular binding site for the urokinase-type plasminogen activator. J Biol Chem 1989; 264:1180-1189) to enrich membrane proteins. The detergent phase was collected and diluted in buffer containing 1% Empigen BB (Calbiochem) to prevent reaggregation of the Triton X-114.

[0171] This material was subjected to Concanavalin A agarose chromatography (Rettig WJ, Garin-Chesa P, Healey JH, Su SL, Ozer HL, Schwab, M, Albino AP, Old LJ. Regulation and heteromeric structure of the fibroblast activation protein in normal and transformed cells of mesenchymal and neuroectodermal origin. Cancer Res 1993; 53:3327-3335).

Biotinylation of L_AH_C

[0172] L_AH_C (1-2 mg) was dialyzed against 50mM bicarbonate buffer and biotinylated with a ten-fold molar excess of

sulfosuccinimidyl-6-biotinamido hexanoate (NHS-LC biotin, Pierce Chemical, Rockford, Illinois, USA) for 2 hours at room temperature. Unreacted product was removed by repeated microdialysis in a microconcentrator.

Transient transfections

[0173] COS-7 cells (American Type Tissue Culture Collection, reference number CRL 1651) were cotransfected by electroporation with the heavy and light chain vectors encoding L_AH_C.

[0174] Anti-CD8 monoclonal antibody was immobilized onto microtiter plates. CD8-FAP from medium of insect cells infected with CD8-FAP baculovirus was allowed to bind to these plates. Spent medium from COS-7 cell cultures transiently transfected with two separate vectors encoding L_AH_C was serially diluted and added to the wells containing the immobilized CD8-FAP. L_AH_C bound to isolated immobilized CD8-FAP protein (Figure 35). Culture supernatants from mock-transfected COS-7 cells failed to demonstrate binding.

[0175] Recombinant membrane-bound FAP from detergent extracts of 293FAP I/2 cells or control extracts was serially diluted and immobilized via chimeric F19 monoclonal antibody bound to microtiter plates. Biotinylated L_AH_C bound recombinant human FAP immobilized with cF19 (Figure 36) in a concentration-dependent manner.

[0176] L_AH_C recognized isolated immobilized recombinant human FAP carrying the epitope for murine F19. L_AH_C bound to both CD8-FAP produced in insect cells, as well as FAP protein produced in 293FAP I/2 cells.

[0177] Culture supernatants from COS7 cells transfected with either heavy and light chain vectors encoding L_AH_C or without DNA (Control) were collected three days posttransfection. CD8-FAP was immobilized via an anti-CD8 antibody as described in the text. Serial dilutions of the COS7 supernatants were allowed to bind to the immobilized CD8-FAP and subsequently detected with an HRP-conjugated anti-human IgG1 antibody.

[0178] Detergent extracts of FAP-expressing 293FAP I/2 cells or control 293 cells were serially diluted and added to cF19-coated microtiter plates. Biotinylated L_AH_C was added and binding of biotinylated L_AH_C was detected with HRP-conjugated streptavidin.

Example 13: Characterization of HT-1080 fibrosarcoma cells and 293 human embryonic kidney cells transfected with cDNA for human FAP

[0179] Fibroblast activation protein (FAP) is a cell-surface, membrane-bound protein which carries the F19 epitope and is expressed on tumor stromal fibroblasts. Cell lines expressing recombinant FAP protein and matched controls lacking FAP were generated for the characterization of anti-FAP monoclonal antibodies.

[0180] Cells used were HT-1080 cells (reference number CCL 121) and 293 human embryonic kidney cells (reference number CRL 1573) were obtained from the American Type Culture Collection (Maryland, USA). Transfectam was obtained from Promega. Geneticin and all restriction enzymes were obtained from Boehringer Mannheim. DNA for transfections was purified from E. coli cells using QiaFilter Maxi Cartridges (Qiagen) as directed by the manufacturer. All DNA preparations were examined by restriction enzyme digestion. Vector sequences were confirmed using an ABI PRISM 310 Sequencer.

[0181] Further information regarding the vectors and DNA sequences employed has been described in Scanlan MJ, Raj BK, Calvo B, Garin-Chesa P, Sanz-Moncasi MP, Healey JH, Old LJ, Rettig WJ. Molecular cloning of fibroblast activation protein alpha, a member of the serine protease family selectively expressed in stromal fibroblasts of epithelial cancers. Proc Natl Acad Sci USA 1992; 89:10832-10836. The FAP cDNA sequence has been deposited in Genbank (accession number HS09287).

Cell culture and immunoassays

[0182] HT-1080 cells were transfected with 1 mg DNA using Transfectam according to the manufacturer's instructions. Human embryonic kidney 293 cells were transfected by calcium phosphate transfection (Brann MR; Buckley NJ; Jones SVP; Bonner TI).

[0183] Expression of cloned muscarinic receptor in A9 L cells. Mol Pharmacol 1987; 32:450-455) with 10 mg DNA. Twenty-four hours later, cells were diluted 1:10 into fresh medium containing 200 mg/mL Geneticin. Colonies were picked and examined by immunofluorescence for FAP expression as described in Rettig WJ; Garin-Chesa P; Beresford HR; Oettgen HF; Melamed MR; Old LJ. Cell-surface glycoproteins of human sarcomas: differential expression in normal and malignant tissues and cultured cells. Proc Natl Acad Sci USA 1988; 85:3110-3114.

[0184] Immunoprecipitations with cF19 were performed with metabolically labelled cells as described in Rettig WJ, Garin-Chesa P, Healey JH, Su SL, Ozer HL, Schwab, M, Albino AP, Old LJ. Regulation and heteromeric structure of the fibroblast activation protein in normal and transformed cells of mesenchymal and neuroectodermal origin. Cancer Res 1993; 53:3327-3335.

[0185] HT-1080 and 293 cells were tested for FAP antigen expression in immunofluorescence assays with anti-FAP

antibodies and were found to be antigen-negative. Transfection of these cells with FAP.38 vector resulted in the generation of Geneticin-resistant colonies. Isolated colonies were picked and analyzed by immunofluorescence for FAP expression. Two cell clones were identified, designated HT-1080FAP clone 33 and 293FAP I/2, which express cell surface-bound FAP protein, as recognized by cF19 antibody. Staining of nonpermeabilized HT-1080FAP clone 33 cells and 293FAP I/2 with cF19 antibody confirmed the cell surface localization of the FAP protein.

[0186] Immunoprecipitation of radiolabelled FAP protein with cF19 from extracts of ³⁵S-methionine labelled HT-1080FAP clone 33 cells or 293FAP I/2 cells resulted in the appearance of a 93 kilodalton band after autoradiography. This band is absent in immunoprecipitates of parental HT-1080 or 293 cell extracts.

[0187] Two stably transfected cell lines, HT-1080FAP clone 33 and 293FAP I/2, express FAP on the cell surface as determined in immunological assays with anti-FAP mAbs. Neither parental HT-1080 cells nor parental 293 cells express detectable levels of FAP.

Example 14: Generation and characterization of CD8-FAP fusion protein

[0188] A soluble form of human FAP (fibroblast activation protein) in the form of a CD8-FAP fusion protein was produced in insect cells for the characterization of L_AH_C containing the binding site for anti-FAP mAbs. Murine CD8 was chosen to permit secretion of the protein and to provide an additional epitope tag.

[0189] The cDNA encoding the extracellular domain of CD8, consisting of the first 189 amino acids of murine CD8, was linked to that of the extracellular domain of FAP (amino acids 27 to 760), essentially as described by Lane, et al. (Lane P, Broucker T, Hubele S, Padovan E, Lazavacchia A, McConnell. Soluble CD40 ligand can replace the normal T cell-derived CD40 ligand signal to B cells in T cell-dependent activation. J Exp Med 1993, 177:1209-1213) using standard PCR protocols. The authenticity of all clones was verified by DNA sequencing. The resulting DNA was inserted into the pVL1393 vector (Invitrogen) and transfection of Sf9 cells (Invitrogen) with this vector and amplification of the resulting recombinant baculovirus were performed as described (Baculovirus Expression Vectors. A Laboratory Manual. O'Reilly DR, Miller LK, Luckow VA, (Eds.), Oxford University Press: New York, 1994). The spent medium of High Five™ cells (Invitrogen) infected with recombinant CD8-FAP baculovirus for four days was collected and cleared by ultracentrifugation.

[0190] The CD8-FAP ELISA (enzyme-linked immunosorbent assay) has been described above (Example 12).

[0191] Insect cell cultures infected with CD8-FAP virus secreted a fusion protein into the medium which carries the F19 epitope and is recognized by an anti-FAP antibody (Figure 1). Neither the cell culture medium alone nor medium from insect cells infected with CD8-CD40L fusion protein bound anti-FAP antibody.

[0192] Soluble CD8-FAP protein carrying the F19 epitope was secreted into the medium of infected insected cell cultures. Culture supernatant from cells infected with a control construct did not contain antigen bearing the F19 epitope.

[0193] A soluble form of FAP, CD8-FAP, was produced in insect cells and CD8-FAP was shown to carry the epitope recognized by cF19.

[0194] Supernatants from insect cells infected with recombinant baculovirus encoding either CD8-FAP or CD8-CD40L fusion protein were collected four days postinfection. Cell culture medium without cells was used as an additional control (medium). Serial dilutions of these materials were added to anti-CD8 antibody-coated microtiter plates and allowed to bind. cF19 (1 mg/mL) was subsequently added and allowed to bind.

[0195] Bound cF19 was detected with horseradish peroxidase-conjugated anti-human antibody.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT:

(A) NAME: Boehringer Ingelheim International GmbH
 (B) STREET: Rheinstrasse
 (C) CITY: Ingelheim am Rhein
 (E) COUNTRY: Germany
 (F) POSTAL CODE (ZIP): 55216
 (G) TELEPHONE: ++49-6132-772770
 (H) TELEFAX: ++49-6132-774377

(ii) TITLE OF INVENTION: FAP alpha-specific antibody with improved
 producibility

(iii) NUMBER OF SEQUENCES: 101

(iv) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk
 (B) COMPUTER: IBM PC compatible
 (C) OPERATING SYSTEM: PC-DOS/MS-DOS
 (D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 339 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

GACATTGTGA	TGACCCAATC	TCCAGACTCT	TTGGCTGTGT	CTCTAGGGGA	GAGGGCCACC	60
ATCAACTGCA	AGTCCAGTCA	GAGCCTTTTA	TATTCTAGAA	ATCAAAAGAA	CTACTTGGCC	120
TGGTATCAGC	AGAAACCAGG	ACAGCCACCC	AAACTCCTCA	TCTTTTGGGC	TAGCACTAGG	180
GAATCTGGGG	TACCTGATAG	GTTCACTGGC	AGTGGGTTTG	GGACAGACTT	CACCCTCACC	240
ATTAGCAGCC	TGCAGGCTGA	AGATGTGGCA	GTTTATTACT	GTCAGCAATA	TTTTAGCTAT	300
CCGCTCACGT	TCGGACAAGG	GACCAAGGTG	GAAATAAAA			339

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 113 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

Asp	Ile	Val	Met	Thr	Gln	Ser	Pro	Asp	Ser	Leu	Ala	Val	Ser	Leu	Gly
1				5					10					15	

Glu Arg Ala Thr Ile Asn Cys Lys Ser Ser Gln Ser Leu Leu Tyr Ser
 20 25 30
 5 Arg Asn Gln Lys Asn Tyr Leu Ala Trp Tyr Gln Gln Lys Pro Gly Gln
 35 40 45
 Pro Pro Lys Leu Leu Ile Phe Trp Ala Ser Thr Arg Glu Ser Gly Val
 50 55 60
 10 Pro Asp Arg Phe Ser Gly Ser Gly Phe Gly Thr Asp Phe Thr Leu Thr
 65 70 75 80
 Ile Ser Ser Leu Gln Ala Glu Asp Val Ala Val Tyr Tyr Cys Gln Gln
 85 90 95
 Tyr Phe Ser Tyr Pro Leu Thr Phe Gly Gln Gly Thr Lys Val Glu Ile
 100 105 110
 15 Lys

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 339 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

GACATTGTGA TGACCCAATC TCCAGACTCT TTGGCTGTGT CTCTAGGGGA GAGGGCCACC 60
 30 ATCAACTGCA AGTCCAGTCA GAGCCTTTTA TATTCTAGAA ATCAAAAGAA CTACTTGGCC 120
 TGGTTCCAGC AGAAACCAGG ACAGCCACCC AAACCTCCTCA TCTTTTGGGC TAGCACTAGG 180
 GAATCTGGGG TACCTGATAG GTTCAGTGGC AGTGGGTTTG GGACAGACTT CACCCTCACC 240
 35 ATTAGCAGCC TGCAGGCTGA AGATGTGGCA GTTTATGACT GTCAACAATA TTTTAGCTAT 300
 CCGCTCACGT TCGGACAAGG GACCAAGGTG GAAATAAAA 339

(2) INFORMATION FOR SEQ ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 113 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

Asp Ile Val Met Thr Gln Ser Pro Asp Ser Leu Ala Val Ser Leu Gly
 1 5 10 15
 50 Glu Arg Ala Thr Ile Asn Cys Lys Ser Ser Gln Ser Leu Leu Tyr Ser
 20 25 30

EP 0 953 639 A1

Arg Asn Gln Lys Asn Tyr Leu Ala Trp Phe Gln Gln Lys Pro Gly Gln
 35 40 45
 Pro Pro Lys Leu Leu Ile Phe Trp Ala Ser Thr Arg Glu Ser Gly Val
 5 50 55 60
 Pro Asp Arg Phe Ser Gly Ser Gly Phe Gly Thr Asp Phe Thr Leu Thr
 65 70 75 80
 Ile Ser Ser Leu Gln Ala Glu Asp Val Ala Val Tyr Asp Cys Gln Gln
 10 85 90 95
 Tyr Phe Ser Tyr Pro Leu Thr Phe Gly Gln Gly Thr Lys Val Glu Ile
 100 105 110
 Lys

(2) INFORMATION FOR SEQ ID NO: 5:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 339 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

GACATTGTGA TGACCCAATC TCCAGACTCT TTGGCTGTGT CTCTAGGGGA GAGGGCCACC 60
 ATCAACTGCA AGTCCAGTCA GAGCCTTTTA TATTCTAGAA ATCAAAAGAA CTACTTGGCC 120
 TGGTATCAGC AGAAACCAGG ACAGCCACCC AAACCTCCTCA TCTATTGGGC TAGCACTAGG 180
 GAATCTGGGG TACCTGATAG GTTCAGTGGC AGTGGGTTTG GGACAGACTT CACCCTCACC 240
 ATTAGCAGCC TGCAGGCTGA AGATGTGGCA GTTTATTACT GTCAGCAATA TTTTAGCTAT 300
 CCGCTCACGT TCGGACAAGG GACCAAGGTG GAAATAAAA 339

(2) INFORMATION FOR SEQ ID NO: 6:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 113 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

Asp Ile Val Met Thr Gln Ser Pro Asp Ser Leu Ala Val Ser Leu Gly
 1 5 10 15
 Glu Arg Ala Thr Ile Asn Cys Lys Ser Ser Gln Ser Leu Leu Tyr Ser
 20 25 30
 Arg Asn Gln Lys Asn Tyr Leu Ala Trp Tyr Gln Gln Lys Pro Gly Gln
 35 40 45
 Pro Pro Lys Leu Leu Ile Tyr Trp Ala Ser Thr Arg Glu Ser Gly Val

EP 0 953 639 A1

50 55 60
 5 Pro Asp Arg Phe Ser Gly Ser Gly Phe Gly Thr Asp Phe Thr Leu Thr
 65 70 75 80
 Ile Ser Ser Leu Gln Ala Glu Asp Val Ala Val Tyr Tyr Cys Gln Gln
 85 90 95
 Tyr Phe Ser Tyr Pro Leu Thr Phe Gly Gln Gly Thr Lys Val Glu Ile
 100 105 110
 10 Lys

(2) INFORMATION FOR SEQ ID NO: 7:

15 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 372 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

CAGGTGCAAC TAGTGCAGTC CGGCGCCGAA GTGAAGAAAC CCGGTGCTTC CGTGAAAGTC 60
 25 AGCTGTAAAA CTAGTAGATA CACCTTCACT GAATACACCA TACACTGGGT TAGACAGGCC 120
 CCTGGCCAAA GGCTGGAGTG GATAGGAGGT ATTAATCCTA ACAATGGTAT TCCTAACTAC 180
 AACCAGAAGT TCAAGGGCCG GGCCACCTTG ACCGTAGGCA AGTCTGCCAG CACCGCCTAC 240
 ATGGAACTGT CCAGCCTGCG CTCCGAGGAC ACTGCAGTCT ACTACTGCGC CAGAAGAAGA 300
 30 ATCGCCTATG GTTACGACGA GGGCCATGCT ATGGACTACT GGGGTCAAGG AACCCCTTGTC 360
 ACCGTCTCCT CA 372

(2) INFORMATION FOR SEQ ID NO: 8:

35 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 124 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

40

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

Gln Val Gln Leu Val Gln Ser Gly Ala Glu Val Lys Lys Pro Gly Ala
 1 5 10 15
 Ser Val Lys Val Ser Cys Lys Thr Ser Arg Tyr Thr Phe Thr Glu Tyr
 20 25 30
 Thr Ile His Trp Val Arg Gln Ala Pro Gly Gln Arg Leu Glu Trp Ile
 35 40 45
 Gly Gly Ile Asn Pro Asn Asn Gly Ile Pro Asn Tyr Asn Gln Lys Phe
 50 55 60

55

Lys Gly Arg Ala Thr Leu Thr Val Gly Lys Ser Ala Ser Thr Ala Tyr
 65 70 75 80
 Met Glu Leu Ser Ser Leu Arg Ser Glu Asp Thr Ala Val Tyr Tyr Cys
 5 85 90 95
 Ala Arg Arg Arg Ile Ala Tyr Gly Tyr Asp Glu Gly His Ala Met Asp
 100 105 110
 Tyr Trp Gly Gln Gly Thr Leu Val Thr Val Ser Ser
 115 120

(2) INFORMATION FOR SEQ ID NO: 9:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 372 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

CAGGTGCAAC TAGTGACAGTC CGGCGCCGAA GTGAAGAAAC CCGGTGCTTC CGTGAAAGTC 60
 AGCTGTAAAA CTAGTAGATA CACCTTCACT GAATACACCA TACACTGGGT TAGACAGGCC 120
 CCTGGCCAAA GGCTGGAGTG GATAGGAGGT ATTAATCCTA ACAATGGTAT TCCTAACTAC 180
 AACCAGAAAGT TCAAGGGCCG GGCCACCTTG ACCGTAGGCA AGTCTGCCAG CACCGCCTAC 240
 ATGGAAGTGT CCAGCCTGCG CTCCGAGGAC ACTGCAGTCT ACTTCTGCGC CAGAAGAAGA 300
 ATCGCCTATG GTTACGACGA GGGCCATGCT ATGGACTACT GGGGTCAAGG AACCCCTTGTC 360
 ACCGTCTCCT CA 372

(2) INFORMATION FOR SEQ ID NO: 10:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 124 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

Gln Val Gln Leu Val Gln Ser Gly Ala Glu Val Lys Lys Pro Gly Ala
 1 5 10 15
 Ser Val Lys Val Ser Cys Lys Thr Ser Arg Tyr Thr Phe Thr Glu Tyr
 20 25 30
 Thr Ile His Trp Val Arg Gln Ala Pro Gly Gln Arg Leu Glu Trp Ile
 35 40 45
 Gly Gly Ile Asn Pro Asn Asn Gly Ile Pro Asn Tyr Asn Gln Lys Phe
 50 55 60
 Lys Gly Arg Ala Thr Leu Thr Val Gly Lys Ser Ala Ser Thr Ala Tyr
 65 70 75 80

EP 0 953 639 A1

Met Glu Leu Ser Ser Leu Arg Ser Glu Asp Thr Ala Val Tyr Phe Cys
85 90 95

Ala Arg Arg Arg Ile Ala Tyr Gly Tyr Asp Glu Gly His Ala Met Asp
100 105 110

Tyr Trp Gly Gln Gly Thr Leu Val Thr Val Ser Ser
115 120

(2) INFORMATION FOR SEQ ID NO: 11:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 372 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

CAGGTGCAAC TAGTGCAATC CGGCGCCGAA GTGAAGAAAC CCGGTGCTTC CGTGAAAGTC	60
AGCTGTAAAA CTAGTAGATA CACCTTCACT GAATACACCA TACACTGGGT TAGACAGGCC	120
CCTGGCCAAA GGCTGGAGTG GATAGGAGGT ATTAATCCTA ACAATGGTAT TCCTAACTAC	180
AACCAGAAGT TCAAGGGCCG GGTCACCATC ACCGTAGACA CCTCTGCCAG CACCGCCTAC	240
ATGGAAGTGT CCAGCCTGCG CTCCGAGGAC ACTGCACTCT ACTACTGCGC CAGAAGAAGA	300
ATCGCCTATG GTTACGACGA GGGCCATGCT ATGGACTACT GGGGTCAAGG AACCCCTTGT	360
ACCGTCTCTCT CA	372

(2) INFORMATION FOR SEQ ID NO: 12:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 124 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:

Gln Val Gln Leu Val Gln Ser Gly Ala Glu Val Lys Lys Pro Gly Ala	1 5 10 15
Ser Val Lys Val Ser Cys Lys Thr Ser Arg Tyr Thr Phe Thr Glu Tyr	20 25 30
Thr Ile His Trp Val Arg Gln Ala Pro Gly Gln Arg Leu Glu Trp Ile	35 40 45
Gly Gly Ile Asn Pro Asn Asn Gly Ile Pro Asn Tyr Asn Gln Lys Phe	50 55 60
Lys Gly Arg Val Thr Ile Thr Val Asp Thr Ser Ala Ser Thr Ala Tyr	65 70 75 80
Met Glu Leu Ser Ser Leu Arg Ser Glu Asp Thr Ala Val Tyr Tyr Cys	

EP 0 953 639 A1

85 90 95

Ala Arg Arg Arg Ile Ala Tyr Gly Tyr Asp Glu Gly His Ala Met Asp
100 105 110

5 Tyr Trp Gly Gln Gly Thr Leu Val Thr Val Ser Ser
115 120

(2) INFORMATION FOR SEQ ID NO: 13:

10 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 372 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

15 (ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:

20 CAGGTGCAAC TAGTGCACTC CGGCGCCGAA GTGAAGAAAC CCGGTGCTTC CGTGAAAGTC 60
AGCTGTAAAA CTAGTAGATA CACCTTCACT GAATACACCA TACACTGGGT TAGACAGGCC 120
CCTGGCCAAA GGCTGGAGTG GATAGGAGGT ATTAATCCTA ACAATGGTAT TCCTAACTAC 180
AACCAGAAGT TCAAGGGCCG GGTCAACATC ACCGTAGACA CCTCTGCCAG CACCGCCTAC 240
25 ATGGAAGTGT CCAGCCTGCG CTCCGAGGAC ACTGCAGTCT ACTTCTGCGC CAGAAGAAGA 300
ATCGCCTATG GTTACGACGA GGGCCATGCT ATGGACTACT GGGGTCAAGG AACCCCTGTC 360
ACCGTCTCCT CA 372

(2) INFORMATION FOR SEQ ID NO: 14:

30 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 124 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

35 (ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:

40 Gln Val Gln Leu Val Gln Ser Gly Ala Glu Val Lys Lys Pro Gly Ala
1 5 10 15
Ser Val Lys Val Ser Cys Lys Thr Ser Arg Tyr Thr Phe Thr Glu Tyr
20 25 30
45 Thr Ile His Trp Val Arg Gln Ala Pro Gly Gln Arg Leu Glu Trp Ile
35 40 45
Gly Gly Ile Asn Pro Asn Asn Gly Ile Pro Asn Tyr Asn Gln Lys Phe
50 55 60
50 Lys Gly Arg Val Thr Ile Thr Val Asp Thr Ser Ala Ser Thr Ala Tyr
65 70 75 80
Met Glu Leu Ser Ser Leu Arg Ser Glu Asp Thr Ala Val Tyr Phe Cys
85 90 95

55

EP 0 953 639 A1

Ala Arg Arg Arg Ile Ala Tyr Gly Tyr Asp Glu Gly His Ala Met Asp
100 105 110

Tyr Trp Gly Gln Gly Thr Leu Val Thr Val Ser Ser
115 120

(2) INFORMATION FOR SEQ ID NO: 15:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 372 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15:

CAGGTGCAAC TAGTGCAGTC CGGCGCCGAA GTGAAGAAAC CCGGTGCTTC CGTGAAAGTC 60
AGCTGTAAAA CTAGTGGATA CACCTTCACT GAATACACCA TACACTGGGT TAGACAGGCC 120
CCTGGCCAAA GGCTGGAGTG GATAGGAGGT ATTAATCCTA ACAATGGTAT TCCTAACTAC 180
AACCAGAAGT TCAAGGGCCG GGTCACCATC ACCGTAGACA CCTCTGCCAG CACCGCCTAC 240
ATGGAAGTGT CCAGCCTGCG CTCCGAGGAC ACTGCAGTCT ACTACTGCGC CAGAAGAAGA 300
ATCGCCTATG GTTACGACGA GGGCCATGCT ATGGACTACT GGGGTCAAGG AACCCCTTGTC 360
ACCGTCTCCT CA 372

(2) INFORMATION FOR SEQ ID NO: 16:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 124 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 16:

Gln Val Gln Leu Val Gln Ser Gly Ala Glu Val Lys Lys Pro Gly Ala
1 5 10 15
Ser Val Lys Val Ser Cys Lys Thr Ser Gly Tyr Thr Phe Thr Glu Tyr
20 25 30
Thr Ile His Trp Val Arg Gln Ala Pro Gly Gln Arg Leu Glu Trp Ile
35 40 45
Gly Gly Ile Asn Pro Asn Asn Gly Ile Pro Asn Tyr Asn Gln Lys Phe
50 55 60
Lys Gly Arg Val Thr Ile Thr Val Asp Thr Ser Ala Ser Thr Ala Tyr
65 70 75 80
Met Glu Leu Ser Ser Leu Arg Ser Glu Asp Thr Ala Val Tyr Tyr Cys
85 90 95
Ala Arg Arg Arg Ile Ala Tyr Gly Tyr Asp Glu Gly His Ala Met Asp
100 105 110

Tyr Trp Gly Gln Gly Thr Leu Val Thr Val Ser Ser
115 120

(2) INFORMATION FOR SEQ ID NO: 17:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 220 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 17:

Asp Ile Val Met Ser Gln Ser Pro Ser Ser Leu Ala Val Ser Val Gly
1 5 10 15
 Glu Lys Val Thr Met Ser Cys Lys Ser Ser Gln Ser Leu Leu Tyr Ser
20 25 30
 Arg Asn Gln Lys Asn Tyr Leu Ala Trp Phe Gln Gln Lys Pro Gly Gln
35 40 45
 Ser Pro Lys Leu Leu Ile Phe Trp Ala Ser Thr Arg Glu Ser Gly Val
50 55 60
 Pro Asp Arg Phe Thr Gly Ser Gly Phe Gly Thr Asp Phe Asn Leu Thr
65 70 75 80
 Ile Ser Ser Val Gln Ala Glu Asp Leu Ala Val Tyr Asp Cys Gln Gln
85 90 95
 Tyr Phe Ser Tyr Pro Leu Thr Phe Gly Ala Gly Thr Lys Leu Glu Leu
100 105 110
 Lys Arg Thr Val Ala Ala Pro Ser Val Phe Ile Phe Pro Pro Ser Asp
115 120 125
 Glu Gln Leu Lys Ser Gly Thr Ala Ser Val Val Cys Leu Leu Asn Asn
130 135 140
 Phe Tyr Pro Arg Glu Ala Lys Val Gln Trp Lys Val Asp Asn Ala Leu
145 150 155 160
 Gln Ser Gly Asn Ser Gln Glu Ser Val Thr Glu Gln Asp Ser Lys Asp
165 170 175
 Ser Thr Tyr Ser Leu Ser Ser Thr Leu Thr Leu Ser Lys Ala Asp Tyr
180 185 190
 Glu Lys His Lys Val Tyr Ala Cys Glu Val Thr His Gln Gly Leu Ser
195 200 205
 Ser Pro Val Thr Lys Ser Phe Asn Arg Gly Glu Cys
210 215 220

(2) INFORMATION FOR SEQ ID NO: 18:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 453 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

EP 0 953 639 A1

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 18:

5	Val	Gln	Leu	Gln	Gln	Ser	Gly	Pro	Glu	Leu	Val	Lys	Pro	Gly	Ala	Ser	1	5	10	15
	Val	Lys	Met	Ser	Cys	Lys	Thr	Ser	Arg	Tyr	Thr	Phe	Thr	Glu	Tyr	Thr	20	25	30	
	Ile	His	Trp	Val	Arg	Gln	Ser	His	Gly	Lys	Ser	Leu	Glu	Trp	Ile	Gly	35	40	45	
10	Gly	Ile	Asn	Pro	Asn	Asn	Gly	Ile	Pro	Asn	Tyr	Asn	Gln	Lys	Phe	Lys	50	55	60	
	Gly	Arg	Ala	Thr	Leu	Thr	Val	Gly	Lys	Ser	Ser	Ser	Thr	Ala	Tyr	Met	65	70	75	80
15	Glu	Leu	Arg	Ser	Leu	Thr	Ser	Glu	Asp	Ser	Ala	Val	Tyr	Phe	Cys	Ala	85	90	95	
	Arg	Arg	Arg	Ile	Ala	Tyr	Gly	Tyr	Asp	Glu	Gly	His	Ala	Met	Asp	Tyr	100	105	110	
20	Trp	Gly	Gln	Gly	Thr	Ser	Val	Thr	Val	Ser	Ser	Ala	Ser	Thr	Lys	Gly	115	120	125	
	Pro	Ser	Val	Phe	Pro	Leu	Ala	Pro	Ser	Ser	Lys	Ser	Thr	Ser	Gly	Gly	130	135	140	
25	Thr	Ala	Ala	Leu	Gly	Cys	Leu	Val	Lys	Asp	Tyr	Phe	Pro	Glu	Pro	Val	145	150	155	160
	Thr	Val	Ser	Trp	Asn	Ser	Gly	Ala	Leu	Thr	Ser	Gly	Val	His	Thr	Phe	165	170	175	
30	Pro	Ala	Val	Leu	Gln	Ser	Ser	Gly	Leu	Tyr	Ser	Leu	Ser	Ser	Val	Val	180	185	190	
	Thr	Val	Pro	Ser	Ser	Ser	Leu	Gly	Thr	Gln	Thr	Tyr	Ile	Cys	Asn	Val	195	200	205	
35	Asn	His	Lys	Pro	Ser	Asn	Thr	Lys	Val	Asp	Lys	Lys	Val	Glu	Pro	Lys	210	215	220	
	Ser	Cys	Asp	Lys	Thr	His	Thr	Cys	Pro	Pro	Cys	Pro	Ala	Pro	Glu	Leu	225	230	235	240
40	Leu	Gly	Gly	Pro	Ser	Val	Phe	Leu	Phe	Pro	Pro	Lys	Pro	Lys	Asp	Thr	245	250	255	
	Leu	Met	Ile	Ser	Arg	Thr	Pro	Glu	Val	Thr	Cys	Val	Val	Val	Asp	Val	260	265	270	
45	Ser	His	Glu	Asp	Pro	Glu	Val	Lys	Phe	Asn	Trp	Tyr	Val	Asp	Gly	Val	275	280	285	
	Glu	Val	His	Asn	Ala	Lys	Thr	Lys	Pro	Arg	Glu	Glu	Gln	Tyr	Asn	Ser	290	295	300	
50	Thr	Tyr	Arg	Val	Val	Ser	Val	Leu	Thr	Val	Leu	His	Gln	Asp	Trp	Leu	305	310	315	320
	Asn	Gly	Lys	Glu	Tyr	Lys	Cys	Lys	Val	Ser	Asn	Lys	Ala	Leu	Pro	Ala	325	330	335	
55	Pro	Ile	Glu	Lys	Thr	Ile	Ser	Lys	Ala	Lys	Gly	Gln	Pro	Arg	Glu	Pro				

EP 0 953 639 A1

340 345 350
 Gln Val Tyr Thr Leu Pro Pro Ser Arg Glu Glu Met Thr Lys Asn Gln
 355 360 365
 5 Val Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala
 370 375 380
 Val Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr
 385 390 395 400
 10 Pro Pro Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu
 405 410 415
 Thr Val Asp Lys Ser Arg Trp Gln Gln Gly Asn Val Phe Ser Cys Ser
 420 425 430
 15 Val Met His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser
 435 440 445
 Leu Ser Pro Gly Lys
 450

(2) INFORMATION FOR SEQ ID NO: 19:

20 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 321 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

25 (ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 19:

30 CGTACTGTGG CTGCACCATC TGTCTTCATC TTCCCGCCAT CTGATGAGCA GTTGAAATCT 60
 GGAAGTGCCT CTGTTGTGTG CCTGCTGAAT AACTTCTATC CCAGAGAGGC CAAAGTACAG 120
 TGGAAGGTGG ATAACGCCCT CCAATCGGGT AACTCCCAGG AGAGTGTACAG AGAGCAGGAC 180
 AGCAAGGACA GCACCTACAG CCTCAGCAGC ACCCTGACGC TGAGCAAAGC AGACTACGAG 240
 35 AAACACAAAG TCTACGCCTG CGAAGTCACC CATCAGGGCC TGAGCTCGCC CGTCACAAAG 300
 AGCTTCAACA GGGGAGAGTG T 321

(2) INFORMATION FOR SEQ ID NO: 20:

40 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 107 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

45 (ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 20:

50 Arg Thr Val Ala Ala Pro Ser Val Phe Ile Phe Pro Pro Ser Asp Glu
 1 5 10 15
 Gln Leu Lys Ser Gly Thr Ala Ser Val Val Cys Leu Leu Asn Asn Phe
 20 25 30

55

EP 0 953 639 A1

Tyr Pro Arg Glu Ala Lys Val Gln Trp Lys Val Asp Asn Ala Leu Gln
 35 40 45
 Ser Gly Asn Ser Gln Glu Ser Val Thr Glu Gln Asp Ser Lys Asp Ser
 50 55 60
 Thr Tyr Ser Leu Ser Ser Thr Leu Thr Leu Ser Lys Ala Asp Tyr Glu
 65 70 75 80
 Lys His Lys Val Tyr Ala Cys Glu Val Thr His Gln Gly Leu Ser Ser
 85 90 95
 Pro Val Thr Lys Ser Phe Asn Arg Gly Glu Cys
 100 105

(2) INFORMATION FOR SEQ ID NO: 21:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 990 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 21:

GCCTCCACCA AGGGCCCATC GGTCTTCCCC CTGGCACCCCT CCTCCAAGAG CACCTCTGGG 60
 GGCACAGCGG CCTGGGCTG CCTGGTCAAG GACTACTTCC CCGAACCGGT GACGGTGTCTG 120
 TGGAACTCAG GCGCCCTGAC CAGCGGCGTG CACACCTTCC CGGCTGTCCT ACAGTCCTCA 180
 GGACTCTACT CCCTCAGCAG CGTGGTGACC GTGCCCTCCA GCAGCTTGGG CACCCAGACC 240
 TACATCTGCA ACGTGAATCA CAAGCCCAGC AACACCAAGG TGGACAAGAA AGTTGAGCCC 300
 AAATCTTGTTG ACAAACCTCA CACATGCCCA CCGTGCCAG CACCTGAACT CCTGGGGGGA 360
 CCGTCAGTCT TCCTCTTCCC CCCAAAACCC AAGGACACCC TCATGATCTC CCGGACCCCT 420
 GAGGTCACAT GCGTGGTGGT GGACGTGAGC CACGAAGACC CTGAGGTCAA GTTCAACTGG 480
 TACGTGGACG GCGTGGAGGT GCATAATGCC AAGACAAAGC CGCGGGAGGA GCAGTACAAC 540
 AGCACGTACC GGGTGGTCAG CGTCCTCACC GTCTTGACACC AGGACTGGCT GAATGGCAAG 600
 GAGTACAAGT GCAAGGTCTC CAACAAAGCC CTCCCAGCCC CCATCGAGAA AACCATCTCC 660
 AAAGCCAAAG GGCAGCCCCG AGAACCACAG GTGTACACCC TGCCCCCATC CCGGGAGGAG 720
 ATGACCAAGA ACCAGGTCAG CCTGACCTGC CTGGTCAAAG GCTTCTATCC CAGCGACATC 780
 GCCGTGGAGT GGGAGAGCAA TGGGCAGCCG GAGAACAAC ACAAAGACCAC GCCTCCCGTG 840
 CTGGACTCCG ACGGCTCCTT CTTCTCTAC AGCAAGCTCA CCGTGGACAA GAGCAGGTGG 900
 CAGCAGGGGA ACGTCTTCTC ATGCTCCGTG ATGCATGAGG CTCTGCACAA CCACTACACG 960
 CAGAAGAGCC TCTCCCTGTC TCCGGGTAAA 990

(2) INFORMATION FOR SEQ ID NO: 22:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 330 amino acids
 (B) TYPE: amino acid

(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 22:

10	Ala	Ser	Thr	Lys	Gly	Pro	Ser	Val	Phe	Pro	Leu	Ala	Pro	Ser	Ser	Lys	1	5	10	15
	Ser	Thr	Ser	Gly	Gly	Thr	Ala	Ala	Leu	Gly	Cys	Leu	Val	Lys	Asp	Tyr	20	25	30	
15	Phe	Pro	Glu	Pro	Val	Thr	Val	Ser	Trp	Asn	Ser	Gly	Ala	Leu	Thr	Ser	35	40	45	
	Gly	Val	His	Thr	Phe	Pro	Ala	Val	Leu	Gln	Ser	Ser	Gly	Leu	Tyr	Ser	50	55	60	
20	Leu	Ser	Ser	Val	Val	Thr	Val	Pro	Ser	Ser	Ser	Leu	Gly	Thr	Gln	Thr	65	70	75	80
	Tyr	Ile	Cys	Asn	Val	Asn	His	Lys	Pro	Ser	Asn	Thr	Lys	Val	Asp	Lys	85	90	95	
25	Lys	Val	Glu	Pro	Lys	Ser	Cys	Asp	Lys	Thr	His	Thr	Cys	Pro	Pro	Cys	100	105	110	
	Pro	Ala	Pro	Glu	Leu	Leu	Gly	Gly	Pro	Ser	Val	Phe	Leu	Phe	Pro	Pro	115	120	125	
30	Lys	Pro	Lys	Asp	Thr	Leu	Met	Ile	Ser	Arg	Thr	Pro	Glu	Val	Thr	Cys	130	135	140	
	Val	Val	Val	Asp	Val	Ser	His	Glu	Asp	Pro	Glu	Val	Lys	Phe	Asn	Trp	145	150	155	160
35	Tyr	Val	Asp	Gly	Val	Glu	Val	His	Asn	Ala	Lys	Thr	Lys	Pro	Arg	Glu	165	170	175	
	Glu	Gln	Tyr	Asn	Ser	Thr	Tyr	Arg	Val	Val	Ser	Val	Leu	Thr	Val	Leu	180	185	190	
40	His	Gln	Asp	Trp	Leu	Asn	Gly	Lys	Glu	Tyr	Lys	Cys	Lys	Val	Ser	Asn	195	200	205	
	Lys	Ala	Leu	Pro	Ala	Pro	Ile	Glu	Lys	Thr	Ile	Ser	Lys	Ala	Lys	Gly	210	215	220	
45	Gln	Pro	Arg	Glu	Pro	Gln	Val	Tyr	Thr	Leu	Pro	Pro	Ser	Arg	Glu	Glu	225	230	235	240
	Met	Thr	Lys	Asn	Gln	Val	Ser	Leu	Thr	Cys	Leu	Val	Lys	Gly	Phe	Tyr	245	250	255	
50	Pro	Ser	Asp	Ile	Ala	Val	Glu	Trp	Glu	Ser	Asn	Gly	Gln	Pro	Glu	Asn	260	265	270	
	Asn	Tyr	Lys	Thr	Thr	Pro	Pro	Val	Leu	Asp	Ser	Asp	Gly	Ser	Phe	Phe	275	280	285	
55	Leu	Tyr	Ser	Lys	Leu	Thr	Val	Asp	Lys	Ser	Arg	Trp	Gln	Gln	Gly	Asn	290	295	300	
	Val	Phe	Ser	Cys	Ser	Val	Met	His	Glu	Ala	Leu	His	Asn	His	Tyr	Thr	305	310	315	320

Gln Lys Ser Leu Ser Leu Ser Pro Gly Lys
325 330

(2) INFORMATION FOR SEQ ID NO: 23:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 427 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 23:

AAGCTTGCCG CCACCATGGA TTCACAGGCC CAGGTTCTTA TGTTACTGCC GCTATGGGTA 60
 TCTGGTACCT GTGGGGACAT TGTGATGTCA CAGTCTCCAT CCTCCCTAGC TGTGTCAGTT 120
 GGAGAGAAGG TTACTATGAG CTGCAAGTCC AGTCAGAGCC TTTTATATAG TCGTAATCAA 180
 AAGAACTACT TGGCCTGGTT CCAGCAGAAG CCAGGGCAGT CTCCTAAACT GCTGATTTTC 240
 TGGGCATCCA CTAGGGAATC TGGGGTCCCT GATCGCTTCA CAGGCAGTGG ATTTGGGACG 300
 GATTTCAATC TCACCATCAG CAGTGTGCAG GCTGAGGACC TGGCAGTTTA TGA CTGT CAG 360
 CAATATTTTA GCTATCCGCT CACGTTTCGGT GCTGGGACCA AGCTGGAGCT GAAACGTGAG 420
 TGGATCC 427

(2) INFORMATION FOR SEQ ID NO: 24:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 133 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 24:

Met Asp Ser Gln Ala Gln Val Leu Met Leu Leu Pro Leu Trp Val Ser
1 5 10
 Gly Thr Cys Gly Asp Ile Val Met Ser Gln Ser Pro Ser Ser Leu Ala
20 25 30
 Val Ser Val Gly Glu Lys Val Thr Met Ser Cys Lys Ser Ser Gln Ser
35 40 45
 Leu Leu Tyr Ser Arg Asn Gln Lys Asn Tyr Leu Ala Trp Phe Gln Gln
50 55 60
 Lys Pro Gly Gln Ser Pro Lys Leu Leu Ile Phe Trp Ala Ser Thr Arg
65 70 75 80
 Glu Ser Gly Val Pro Asp Arg Phe Thr Gly Ser Gly Phe Gly Thr Asp
85 90 95
 Phe Asn Leu Thr Ile Ser Ser Val Gln Ala Glu Asp Leu Ala Val Tyr
100 105 110

Asp Cys Gln Gln Tyr Phe Ser Tyr Pro Leu Thr Phe Gly Ala Gly Thr
 115 120 125

Lys Leu Glu Leu Lys
 130

(2) INFORMATION FOR SEQ ID NO: 25:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 457 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 25:

AAGCTTGCCG CCACCATGGG ATGGAGCTGG GTCTTTCTCT TTCTCCTGTC AGGAACTGCA 60
 GGTGTCCTCT CTGAGGTCCA GCTGCAACAG TCTGGACCTG AGCTGGTGAA GCCTGGGGCT 120
 TCAGTAAAGA TGTCCTGCAA GACTTCTAGA TACACATTCA CTGAATACAC CATACTGG 180
 GTGAGACAGA GCCATGGAAA GAGCCTTGAG TGGATTGGAG GTATTAATCC TAACAATGGT 240
 ATTCCTAACT ACAACCAGAA GTTCAAGGGC AGGGCCACAT TGACTGTAGG CAAGTCCTCC 300
 AGCACCGCCT ACATGGAGCT CCGCAGCCTG ACATCTGAGG ATTCTGCGGT CTATTTCTGT 360
 GCAAGAAGAA GAATCGCCTA TGGTTACGAC GAGGGCCATG CTATGGACTA CTGGGGTCAA 420
 GGAACCTCAG TCACCGTCTC CTCAGGTGAG TGGATCC 457

(2) INFORMATION FOR SEQ ID NO: 26:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 143 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 26:

Met Gly Trp Ser Trp Val Phe Leu Phe Leu Leu Ser Gly Thr Ala Gly
 1 5 10 15
 Val Leu Ser Glu Val Gln Leu Gln Gln Ser Gly Pro Glu Leu Val Lys
 20 25 30
 Pro Gly Ala Ser Val Lys Met Ser Cys Lys Thr Ser Arg Tyr Thr Phe
 35 40 45
 Thr Glu Tyr Thr Ile His Trp Val Arg Gln Ser His Gly Lys Ser Leu
 50 55 60
 Glu Trp Ile Gly Gly Ile Asn Pro Asn Asn Gly Ile Pro Asn Tyr Asn
 65 70 75 80
 Gln Lys Phe Lys Gly Arg Ala Thr Leu Thr Val Gly Lys Ser Ser Ser
 85 90 95

Thr Ala Tyr Met Glu Leu Arg Ser Leu Thr Ser Glu Asp Ser Ala Val
100 105 110

5 Tyr Phe Cys Ala Arg Arg Arg Ile Ala Tyr Gly Tyr Asp Glu Gly His
115 120 125

Ala Met Asp Tyr Trp Gly Gln Gly Thr Ser Val Thr Val Ser Ser
130 135 140

(2) INFORMATION FOR SEQ ID NO: 27:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 8068 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 27:

20	GAATTCCAGC	ACACTGGCGG	CCGTTACTAG	TTATTAATAG	TAATCAATTA	CGGGGTCATT	60
	AGTTCATAGC	CCATATATGG	AGTTCCGCGT	TACATAACTT	ACGGTAAATG	GCCCGCCTGG	120
	CTGACCGCCC	AACGACCCCC	GCCCATTGAC	GTCAATAATG	ACGTATGTTT	CCATAGTAAC	180
	GCCAATAGGG	ACTTTCCATT	GACGTCAATG	GGTGGAGTAT	TTACGGTAAA	CTGCCCCTT	240
25	GGCAGTACAT	CAAGTGTATC	ATATGCCAAG	TACGCCCCCT	ATTGACGTCA	ATGACGGTAA	300
	ATGGCCCCGC	TGGCATTATG	CCCAGTACAT	GACCTTATGG	GACTTTCCTA	CTTGGCAGTA	360
	CATCTACGTA	TTAGTCATCG	CTATTACCAT	GGTGATGCGG	TTTTGGCAGT	ACATCAATGG	420
30	GCGTGGATAG	CGGTTTGACT	CACGGGGATT	TCCAAGTCTC	CACCCCATTTG	ACGTCAATGG	480
	GAGTTTGTIT	TGGCACCAAA	ATCAACGGGA	CTTTCCAAAA	TGTCGTAACA	ACTCCGCCCC	540
	ATTGACGCAA	ATGGGCGGTA	GGCGTGTACG	GTGGGAGGTC	TATATAAGCA	GAGCTCGTTT	600
	AGTGAACCGT	CAGATCGCCT	GGAGACGCCA	TCCACGCTGT	TTTGACCTCC	ATAGAAGACA	660
35	CCGGGACCGA	TCCAGCCTCC	GCGGCCGGGA	ACGGTGCAAT	GGAACGCGGA	TTCCCCGTGC	720
	CAAGAGTGAC	GTAAGTACCG	CCTATAGAGT	CTATAGGCCC	ACCCCTTTGG	CTTCTTATGC	780
	ATGCTATACT	GTTTTTGGCT	TGGGGTCTAT	ACACCCCGCG	TTCTTCATGT	TATAGGTGAT	840
40	GGTATAGCTT	AGCCTATAGG	TGTGGGTAT	TGACCATTAT	TGACCACTCC	CCTATTGGTG	900
	ACGATACTTT	CCATTACTAA	TCCATAACAT	GGCTCTTTGC	CACAACTCTC	TTTATTGGCT	960
	ATATGCCAAT	ACACTGTCCT	TCAGAGACTG	ACACGGACTC	TGTATTTTTA	CAGGATGGGG	1020
45	TCTCATTTAT	TATTTACAAA	TTACATATA	CAACACCACC	GTCCCCAGTG	CCCGCAGTTT	1080
	TTATTAAACA	TAACGTGGGA	TCTCCACGCG	AATCTCGGGT	ACGTGTTCCG	GACATGGGCT	1140
	CTTCTCCGGT	AGCGGCGGAG	CTTCTACATC	CGAGCCCTGC	TCCCATGCCT	CCAGCGACTC	1200
	ATGGTCGCTC	GGCAGCTCCT	TGCTCCTAAC	AGTGGAGGCC	AGACTTAGGC	ACAGCACGAT	1260
50	GCCCACCACC	ACCAGTGTGC	CGCACAAGGC	CGTGGCGGTA	GGGTATGTGT	CTGAAAATGA	1320
	GCTCGGGGAG	CGGGCTTGCA	CCGCTGACGC	ATTTGGAAGA	CTTAAGGCAG	CGGCAGAAGA	1380

EP 0 953 639 A1

	AGATGCAGGC	AGCTGAGTTG	TTGTGTTCTG	ATAAGAGTCA	GAGGTAAGTC	CCGTTGCGGT	1440
	GCTGTTAACG	GTGGAGGGCA	GTGTAGTCTG	AGCAGTACTC	GTTGCTGCCG	CGCGCGCCAC	1500
5	CAGACATAAT	AGCTGACAGA	CTAACAGACT	GTTCCCTTCC	ATGGGTCTTT	TCTGCAGTCA	1560
	CCGTCTTTGA	CACGCGTCTC	GGGAAGCTTG	CCGCCACCAT	GGATTACACAG	GCCCAGGTTC	1620
	TTATGTTACT	GCCGCTATGG	GTATCTGGTA	CCTGTGGGGA	CATTGTGATG	TCACAGTCTC	1680
10	CATCCTCCCT	AGCTGTGTCA	GTTGGAGAGA	AGGTTACTAT	GAGCTGCAAG	TCCAGTCAGA	1740
	GCCTTTTATA	TTCTAGAAAT	CAAAAGAACT	ACTTGGCCTG	GTTCCAGCAG	AAGCCAGGGC	1800
	AGTCTCCTAA	ACTGCTGATT	TTCTGGGCAT	CCACTAGGGA	ATCTGGGGTC	CCTGATCGCT	1860
	TCACAGGCAG	TGGATTTGGG	ACGGATTTC	ATCTCACCAT	CAGCAGTGTG	CAGGCTGAGG	1920
15	ACCTGGCAGT	TTATGACTGT	CAGCAATATT	TTAGCTATCC	GCTCACGTTT	GGTGCTGGGA	1980
	CCAAGCTGGA	GCTGAAACGT	GAGTGGATCC	ATCTGGGATA	AGCATGCTGT	TTTCTGTCTG	2040
	TCCCTAACAT	GCCCTGTGAT	TATGCGCAAA	CAACACACCC	AAGGGCAGAA	CTTTGTTACT	2100
20	TAAACACCAT	CCTGTTTGCT	TCTTTCCTCA	GGAAGTGTGG	CTGCACCATC	TGTCTTCATC	2160
	TTCCCGCCAT	CTGATGAGCA	GTTGAAATCT	GGAAGTGCCT	CTGTTGTGTG	CCTGCTGAAT	2220
	AACCTCTATC	CCAGAGAGGC	CAAAGTACAG	TGGAAGGTGG	ATAACGCCCT	CCAATCGGGT	2280
25	AACCTCCAGG	AGAGTGTAC	AGAGCAGGAC	AGCAAGGACA	GCACCTACAG	CCTCAGCAGC	2340
	ACCTGACGC	TGAGCAAAGC	AGACTACGAG	AAACACAAAG	TCTACGCCTG	CGAAGTCACC	2400
	CATCAGGGCC	TGAGCTCGCC	CGTCACAAAG	AGCTTCAACA	GGGGAGAGTG	TTAGAGGGAG	2460
	AAGTGCCCCC	ACCTGCTCCT	CAGTTCCAGC	CTGACCCCTT	CCCATCCTTT	GGCCTCTGAC	2520
30	CCTTTTTCCA	CAGGGGACCT	ACCCCTATTG	CGGTCTCCA	GCTCATCTTT	CACCTCACCC	2580
	CCCTCCTCCT	CCTTGGCTTT	AATTATGCTA	ATGTTGGAGG	AGAATGAATA	AATAAAGTGA	2640
	ATCTTTGCAC	CTGTGGTGGA	TCTAATAAAA	GATATTTATT	TTCAATTAGAT	ATGTGTGTTG	2700
35	GTTTTTTGTG	TGCAGTGCCT	CTATCTGGAG	GCCAGGTAGG	GCTGGCCTTG	GGGGAGGGGG	2760
	AGGCCAGAAT	GACTCCAAGA	GCTACAGGAA	GGCAGGTCAG	AGACCCCACT	GGACAAACAG	2820
	TGGCTGGACT	CTGCACCATA	ACACACAATC	AACAGGGGAG	TGAGCTGGAA	ATTTGCTAGC	2880
	GAATTCTTGA	AGACGAAAGG	GCCTCGTGAT	ACGCCATATT	TTATAGGTTA	ATGTCATGAT	2940
40	AATAATGGTT	TCTTAGACGT	CAGGTGGCAC	TTTTTCGGGA	AATGTGCGCG	GAACCCCTAT	3000
	TTGTTTATTT	TTCTAAATAC	ATTCAAATAT	GTATCCGCTC	ATGAGACAAT	AACCCTGATA	3060
	AATGCTTCAA	TAATATTGAA	AAAGGAAGAG	TATGAGTATT	CAACATTTCC	GTGTCGCCCT	3120
45	TATTCCTTTT	TTTGCGGCAT	TTTGCTTCC	TGTTTTTGCT	CACCCAGAAA	CGCTGGTGAA	3180
	AGTAAAAGAT	GCTGAAGATC	AGTTGGGTGC	ACGAGTGGGT	TACATCGAAC	TGGATCTCAA	3240
	CAGCGGTAAG	ATCCTTGAGA	GTTTTCGCCC	CGAAGAACGT	TTTCCAATGA	TGAGCACTTT	3300
50	TAAAGTTCTG	CTATGTGGCG	CGGTATTATC	CCGTGTTGAC	GCCGGGCAAG	AGCAACTCGG	3360
	TCGCCGCATA	CACTATTCTC	AGAATGACTT	GGTTGAGTAC	TCACCAGTCA	CAGAAAAGCA	3420
	TCTTACGGAT	GGCATGACAG	TAAGAGAATT	ATGCAGTGCT	GCCATAACCA	TGAGTGATAA	3480
55							

EP 0 953 639 A1

	CACTGCGGCC	AACTTACTTC	TGACAACGAT	CGGAGGACCG	AAGGAGCTAA	CCGCTTTTTT	3540
	GCACAACATG	GGGGATCATG	TAACTCGCCT	TGATCGTTGG	GAACCGGAGC	TGAATGAAGC	3600
5	CATACCAAAC	GACGAGCGTG	ACACCACGAT	GCCTGCAGCA	ATGGCAACAA	CGTTGCGCAA	3660
	ACTATTAACT	GGCGAACTAC	TTACTCTAGC	TTCCCGGCAA	CAATTAATAG	ACTGGATGGA	3720
	GGCGGATAAA	GTTGCAGGAC	CACTTCTGCG	CTCGGCCCTT	CCGGCTGGCT	GGTTTATTGC	3780
10	TGATAAATCT	GGAGCCGGTG	AGCGTGGGTC	TCGCGGTATC	ATTGCAGCAC	TGGGGCCAGA	3840
	TGGTAAGCCC	TCCCGTATCG	TAGTTATCTA	CACGACGGGG	AGTCAGGCAA	CTATGGATGA	3900
	ACGAAATAGA	CAGATCGCTG	AGATAGGTGC	CTCACTGATT	AAGCATTGGT	AACTGTGAGA	3960
	CCAAGTTTAC	TCATATATAC	TTTAGATTGA	TTTAAAACTT	CATTTTTAAT	TTAAAGGAT	4020
15	CTAGGTGAAG	ATCCTTTTTG	ATAATCTCAT	GACCAAAATC	CCTTAACGTG	AGTTTTTCGT	4080
	CCACTGAGCG	TCAGACCCCG	TAGAAAAGAT	CAAAGGATCT	TCTTGAGATC	CTTTTTTTCT	4140
	GCGCGTAATC	TGCTGCTTGC	AAACAAAAAA	ACCACCGCTA	CCAGCGGTGG	TTTGTTTGCC	4200
20	GGATCAAGAG	CTACCAACTC	TTTTTCCGAA	GGTAACTGGC	TTCAAGCAAG	CGCAGATACC	4260
	AAATACTGTC	CTTCTAGTGT	AGCCGTAGTT	AGGCCACCAC	TTCAAGAACT	CTGTAGCACC	4320
	GCCTACATAC	CTCGCTCTGC	TAATCCTGTT	ACCAGTGGCT	GCTGCCAGTG	GCGATAAGTC	4380
25	GTGTCTTACC	GGGTTGGACT	CAAGACGATA	GTTACCGGAT	AAGGCGCAGC	GGTCGGGCTG	4440
	AACGGGGGGT	TCGTGCACAC	AGCCCAGCTT	GGAGCGAACG	ACCTACACCG	AACTGAGATA	4500
	CCTACAGCGT	GAGCTATGAG	AAAGCGCCAC	GCTTCCCGAA	GGGAGAAAGG	CGGACAGGTA	4560
	TCCGGTAAGC	GGCAGGGTCG	GAACAGGAGA	GCGCACGAGG	GAGCTTCCAG	GGGGAAACGC	4620
30	CTGGTATCTT	TATAGTCCTG	TCGGGTTTCG	CCACCTCTGA	CTTGAGCGTC	GATTTTTGTG	4680
	ATGCTCGTCA	GGGGGGCGGA	GCCTATGGAA	AAACGCCAGC	AACGCGGCCT	TTTTTACGGT	4740
	CCTGGCCTTT	TGCTGGCCTT	TTGCTCACAT	GTTCTTTTCT	GCGTTATCCC	CTGATTCTGT	4800
35	GGATAACCGT	ATTACCGCCT	TTGAGTGAGC	TGATACCGCT	CGCCGCAGCC	GAACGACCGA	4860
	GCGCAGCGAG	TCAGTGAGCG	AGGAAGCGGA	AGAGCGCCTG	ATGCGGTATT	TTCTCCTTAC	4920
	GCATCTGTGC	GGTATTTTAC	ACCGCATATG	GTGCACTCTC	AGTACAATCT	GCTCTGATGC	4980
40	CGCATAGTTA	AGCCAGTATA	CACTCCGCTA	TCGCTACGTG	ACTGGGTCAT	GGCTGCGCCC	5040
	CGACACCCGC	CAACACCCGC	TGACGCGCCC	TGACGGGCTT	GTCTGCTCCC	GGCATCCGCT	5100
	TACAGACAAG	CTGTGACCGT	CTCCGGGAGC	TGCATGTGTC	AGAGGTTTTT	ACCGTCATCA	5160
	CCGAAACGCG	CGAGGCAGCT	GTGGAATGTG	TGTCAGTTAG	GGTGTGGAAA	GTCCCCAGGC	5220
45	TCCCCAGCAG	GCAGAAGTAT	GCAAAGCATG	CATCTCAATT	AGTCAGCAAC	CAGGCTCCCC	5280
	AGCAGGCAGA	AGTATGCAAA	GCATGCATCT	CAATTAGTCA	GCAACCATAG	TCCCGCCCCCT	5340
	AACTCCGCCC	ATCCCGCCCC	TAACTCCGCC	CAGTTCCGCC	CATTCTCCGC	CCCATGGCTG	5400
50	ACTAATTTTT	TTTATTTTATG	CAGAGGCCGA	GGCCGCTCG	GCCTCTGAGC	TATTCCAGAA	5460
	GTAGTGAGGA	GGCTTTTTTTG	GAGGCCTAGG	CTTTTGCAAA	AAGCTAGCTT	CACGCTGCCG	5520
	CAAGCACTCA	GGGCGCAAGG	GCTGTAAAG	GAAGCGGAAC	ACGTAGAAAG	CCAGTCCGCA	5580
55							

	GAAACGGTGC	TGACCCCGGA	TGAATGTCAG	CTACTGGGCT	ATCTGGACAA	GGGAAAACGC	5640
	AAGCGCAAAG	AGAAAGCAGG	TAGCTTGCAG	TGGGCTTACA	TGGCGATAGC	TAGACTGGGC	5700
5	GGTTTTATGG	ACAGCAAGCG	AACCGGAATT	GCCAGCTGGG	GCGCCCTCTG	GTAAGGTTGG	5760
	GAAGCCCTGC	AAAGTAAACT	GGATGGCTTT	CTTGCCGCCA	AGGATCTGAT	GGCGCAGGGG	5820
	ATCAAGATCT	GATCAAGAGA	CAGGATGAGG	ATCGTTTCGC	ATGATTGAAC	AAGATGGATT	5880
10	GCACGCAGGT	TCTCCGGCCG	CTTGGGTGGA	GAGGCTATTG	GGCTATGACT	GGGCACAACA	5940
	GACAATCGGC	TGCTCTGATG	CCGCCGTGTT	CCGGCTGTCA	GCGCAGGGGC	GCCCGGTTCT	6000
	TTTTGTCAAG	ACCGACCTGT	CCGGTGCCCT	GAATGAACTG	CAGGACGAGG	CAGCGCGGCT	6060
	ATCGTGGCTG	GCCACGACGG	GCGTTCCTTG	CGCAGCTGTG	CTCGACGTTG	TCACTGAAGC	6120
15	GGGAAGGGAC	TGGCTGCTAT	TGGGCGAAGT	GCCGGGGCAG	GATCTCCTGT	CATCTCACCT	6180
	TGCTCCTGCC	GAGAAAGTAT	CCATCATGGC	TGATGCAATG	CGGCGGCTGC	ATACGCTTGA	6240
	TCCGGCTACC	TGCCCATTCG	ACCACCAAGC	GAAACATCGC	ATCGAGCGAG	CACGTACTCG	6300
20	GATGGAAGCC	GGTCTTGTCG	ATCAGGATGA	TCTGGACGAA	GAGCATCAGG	GGCTCGCGCC	6360
	AGCCGAACTG	TTCGCCAGGC	TCAAGGCGCG	CATGCCCCGAC	GGCGAGGATC	TCGTCTGTGAC	6420
	CCATGGCGAT	GCCTGCTTGC	CGAATATCAT	GGTGGAAAAT	GGCCGCTTTT	CTGGATTTCAT	6480
25	CGACTGTGGC	CGGCTGGGTG	TGGCGGACCG	CTATCAGGAC	ATAGCGTTGG	CTACCCGTGA	6540
	TATTGCTGAA	GAGCTTGCGG	GCGAATGGGC	TGACCCTTTC	CTCGTGCTTT	ACGGTATCGC	6600
	CGCTCCCGAT	TCGCAGCGCA	TCGCCTTCTA	TCGCCTTCTT	GACGAGTTCT	TCTGAGCGGG	6660
	ACTCTGGGGT	TCGAAATGAC	CGACCAAGCG	ACGCCCAACC	TGCCATCACG	AGATTTCGAT	6720
30	TCCACCGCCG	CCTTCTATGA	AAGGTTGGGC	TTCCGGAATCG	TTTTCCGGGA	CGCCGGCTGG	6780
	ATGATCCTCC	AGCGCGGGGA	TCTCATGCTG	GAGTTCTTCG	CCCACCCCGG	GCTCGATCCC	6840
	CTCGCGAGTT	GGTTCAGCTG	CTGCCTGAGG	CTGGACGACC	TCGCGGAGTT	CTACCGGCAG	6900
35	TGCAAATCCG	TCGGCATCCA	GGAAACCAGC	AGCGGCTATC	CGCGCATCCA	TGCCCCCGAA	6960
	CTGCAGGAGT	GGGGAGGCAC	GATGGCCGCT	TTGGTCCCGG	ATCTTTGTGA	AGGAACCTTA	7020
	CTTCTGTGGT	GTGACATAAT	TGGACAAACT	ACCTACAGAG	ATTTAAAGCT	CTAAGGTAAA	7080
40	TATAAAATTT	TTAAGTGTAT	AATGTGTTAA	ACTACTGATT	CTAATTGTTT	GTGTATTTTA	7140
	GATTCCAACC	TATGGAAGTG	ATGAATGGGA	GCAGTGGTGG	AATGCCTTTA	ATGAGGAAAA	7200
	CCTGTTTTGC	TCAGAAGAAA	TGCCATCTAG	TGATGATGAG	GCTACTGCTG	ACTCTCAACA	7260
	TTCTACTCCT	CCAAAAAAGA	AGAGAAAGGT	AGAAGACCCC	AAGGACTTTC	CTTCAGAATT	7320
45	GCTAAGTTTT	TTGAGTCATG	CTGTGTTTAG	TAATAGAACT	CTTGCTTGCT	TTGCTATTTA	7380
	CACCACAAAG	GAAAAAGCTG	CACTGCTATA	CAAGAAAATT	ATGGAAAAAT	ATTCTGTAAC	7440
	CTTTATAAGT	AGGCATAACA	GTTATAATCA	TAACATACTG	TTTTTTCTTA	CTCCACACAG	7500
50	GCATAGAGTG	TCTGCTATTA	ATAACTATGC	TCAAAAATTG	TGTACCTTTA	GCTTTTTAAT	7560
	TTGTAAAGGG	GTTAATAAGG	AATATTTGAT	GTATAGTGCC	TTGACTAGAG	ATCATAATCA	7620
	GCCATACCAC	ATTTGTAGAG	GTTTTACTTG	CTTTAAAAAA	CCTCCACAC	CTCCCCCTGA	7680
55							

EP 0 953 639 A1

```

ACCTGAAACA TAAAATGAAT GCAATTGTG TTGTTAACTT GTTTATTGCA GCTTATAATG 7740
GTTACAAATA AAGCAATAGC ATCACAAATT TCACAAATAA AGCATTTTTT TCACTGCATT 7800
5 CTAGTTGTGG TTTGTCCAAA CTCATCAATG TATCTTATCA TGTCTGGATC TAATAAAAGA 7860
TATTTATTTT CATTAGATAT GTGTGTTGGT TTTTGTGTG CAGTGCCTCT ATCTGGAGGC 7920
CAGGTAGGGC TGGCCTTGGG GGAGGGGGAG GCCAGAATGA CTCCAAGAGC TACAGGAAGG 7980
10 CAGGTCAGAG ACCCCACTGG ACAAACAGTG GCTGGACTCT GCACCATAAC ACACAATCAA 8040
CAGGGGAGTG AGCTGGAAAT TTGCTAGC 8068

```

(2) INFORMATION FOR SEQ ID NO: 28:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 239 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 28:

```

Asp Ser Gln Ala Gln Val Leu Met Leu Leu Pro Leu Trp Val Ser Gly
1      5      10      15
25 Thr Cys Gly Asp Ile Val Met Ser Gln Ser Pro Ser Ser Leu Ala Val
    20      25      30
    Ser Val Gly Glu Lys Val Thr Met Ser Cys Lys Ser Ser Gln Ser Leu
        35      40      45
30 Leu Tyr Ser Arg Asn Gln Lys Asn Tyr Leu Ala Trp Phe Gln Gln Lys
    50      55      60
    Pro Gly Gln Ser Pro Lys Leu Leu Ile Phe Trp Ala Ser Thr Arg Glu
    65      70      75      80
35 Ser Gly Val Pro Asp Arg Phe Thr Gly Ser Gly Phe Gly Thr Asp Phe
    85      90      95
    Asn Leu Thr Ile Ser Ser Val Gln Ala Glu Asp Leu Ala Val Tyr Asp
        100      105      110
    Cys Gln Gln Tyr Phe Ser Tyr Pro Leu Thr Phe Gly Ala Gly Thr Lys
        115      120      125
40 Leu Glu Leu Lys Arg Thr Val Ala Ala Pro Ser Val Phe Ile Phe Pro
    130      135      140
    Pro Ser Asp Glu Gln Leu Lys Ser Gly Thr Ala Ser Val Val Cys Leu
    145      150      155      160
45 Leu Asn Asn Phe Tyr Pro Arg Glu Ala Lys Val Gln Trp Lys Val Asp
        165      170      175
    Asn Ala Leu Gln Ser Gly Asn Ser Gln Glu Ser Val Thr Glu Gln Asp
        180      185      190
50 Ser Lys Asp Ser Thr Tyr Ser Leu Ser Ser Thr Leu Thr Leu Ser Lys
    195      200      205
    Ala Asp Tyr Glu Lys His Lys Val Tyr Ala Cys Glu Val Thr His Gln
    210      215      220
55

```

Gly Leu Ser Ser Pro Val Thr Lys Ser Phe Asn Arg Gly Glu Cys
 225 230 235

(2) INFORMATION FOR SEQ ID NO: 29:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 7731 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 29:

15	TTGAAGACGA AAGGGCCTCG TGATACGCCT ATTTTATAG GTTAATGTCA TGATAATAAT	60
	GGTTTCTTAG ACGTCAGGTG GCACTTTTCG GGGAAATGTG CGCGGAACCC CTATTTGTTT	120
	ATTTTTCTAA ATACATTCAA ATATGTATCC GCTCATGAGA CAATAACCCT GATAAATGCT	180
20	TCAATAATAT TGAAAAAGGA AGAGTATGAG TATTCAACAT TTCCGTGTCG CCCTTATTCC	240
	CTTTTTTGCG GCATTTTGCC TTCCTGTTTT TGCTCACCCA GAAACGCTGG TGAAAGTAAA	300
	AGATGCTGAA GATCAGTTGG GTGCACGAGT GGGTTACATC GAACTGGATC TCAACAGCGG	360
	TAAGATCCTT GAGAGTTTTT GCCCCGAAGA ACGTTTTCCA ATGATGAGCA CTTTAAAGT	420
25	TCTGCTATGT GGC GCGGTAT TATCCCGTGT TGACGCCGG CAAGAGCAAC TCGGTCGCCG	480
	CATACACTAT TCTCAGAATG ACTTG GTTGA GTACTCACCA GTCACAGAAA AGCATCTTAC	540
	GGATGGCATG ACAGTAAGAG AATTATGCAG TGCTGCCATA ACCATGAGTG ATAACACTGC	600
30	GGCCAACTTA CTCTGACAA CGATCGGAGG ACCGAAGGAG CTAACCGCTT TTTGACAA	660
	CATGGGGGAT CATGTAATC GCCTTGATCG TTGGGAACCG GAGCTGAATG AAGCCATACC	720
	AAACGACGAG CGTGACACCA CGATGCCTGC AGCAATGGCA ACAACGTTGC GCAAATATT	780
35	AACTGGCGAA CTACTTACTC TAGCTTCCCG GCAACAATTA ATAGACTGGA TGGAGGCGGA	840
	TAAAGTTGCA GGACCACTT TCGCTCGGC CCTTCCGGCT GGCTGGTTTA TTGCTGATAA	900
	ATCTGGAGCC GGTGAGCGT GGTCTCGCG TATCATTGCA GCACTGGGGC CAGATGGTAA	960
40	GCCCTCCCGT ATCGTAGTTA TCTACACGAC GGGGAGTCAG GCAACTATGG ATGAACGAAA	1020
	TAGACAGATC GCTGAGATAG GTGCCTCACT GATTAAGCAT TGGTAACTGT CAGACCAAGT	1080
	TTACTCATAT AACTTTTAGA TTGATTTAAA ACTTCATTTT TAATTTAAAA GGATCTAGGT	1140
	GAAGATCCTT TTGATAATC TCATGACCAA AATCCCTTAA CGTGAGTTTT CGTTCCACTG	1200
45	AGCGTCAGAC CCCGTAGAAA AGATCAAAGG ATCTTCTTGA GATCCTTTTT TTCTGCGCGT	1260
	AATCTGCTGC TTGCAAACAA AAAAACCACC GCTACCAGCG GTGGTTTGTT TGCCGGATCA	1320
	AGAGCTACCA ACTCTTTTTT CGAAGGTAAC TGGCTTCAGC AGAGCGCAGA TACCAAATAC	1380
50	TGTCCTTCTA GTGTAGCCGT AGTTAGGCCA CCACTTCAAG AACTCTGTAG CACCGCCTAC	1440
	ATACCTCGCT CTGCTAATCC TGTTACCAGT GGCTGCTGCC AGTGCGGATA AGTCGTGTCT	1500
	TACCGGGTTG GACTCAAGAC GATAGTTACC GGATAAGGCG CAGCGGTCGG GCTGAACGGG	1560

EP 0 953 639 A1

	GGGTTCTGTC	ACACAGCCCA	GCTTGGAGCG	AACGACCTAC	ACCGAACTGA	GATACCTACA	1620
	GCGTGAGCTA	TGAGAAAGCG	CCACGCTTCC	CGAAGGGAGA	AAGGCGGACA	GGTATCCGGT	1680
5	AAGCGGCAGG	GTCGGAACAG	GAGAGCGCAC	GAGGGAGCTT	CCAGGGGGAA	ACGCCTGGTA	1740
	TCTTTATAGT	CCTGTCGGGT	ITCGCCACCT	CTGACTTGAG	CGTCGATTTT	TGTGATGCTC	1800
	GTCAGGGGGG	CGGAGCCTAT	GGAAAAACGC	CAGCAACGCG	GCCTTTTTTAC	GGTTCCTGGC	1860
10	CTTTTGCTGG	CCTTTTGCTC	ACATGTTCTT	TCCTGCGTTA	TCCCCTGATT	CTGTGGATAA	1920
	CCGTATTACC	GCCTTTGAGT	GAGCTGATAC	CGCTCGCCGC	AGCCGAACGA	CCGAGCGCAG	1980
	CGAGTCAGTG	AGCGAGGAAG	CGGAAGAGCG	CCTGATGCGG	TATTTTCTCC	TTACGCATCT	2040
15	GTGCGGTATT	TCACACCGCA	TATGGTGAC	TCTCAGTACA	ATCTGCTCTG	ATGCCGCATA	2100
	GTTAAGCCAG	TATACACTCC	GCTATCGCTA	CGTGACTGGG	TCATGGCTGC	GCCCCGACAC	2160
	CCGCCAACAC	CCGCTGACGC	GCCCTGACGG	GCTTGTCTGC	TCCCGGCATC	CGCTTACAGA	2220
	CAAGCTGTGA	CCGTCTCCGG	GAGCTGCATG	TGTCAGAGGT	TTTACCCGTC	ATCACCGAAA	2280
20	CGCGCGAGGC	AGCATGCATC	TCAATTAGTC	AGCAACCATA	GTCCCGCCCC	TAACCTCCGCC	2340
	CATCCCGCCC	CTAACTCCGC	CCAGTTCCGC	CCATTTCTCG	CCCCATGGCT	GACTAATTTT	2400
	TTTTATTTAT	GCAGAGGCCG	AGGCCGCCTC	GGCCTCTGAG	CTATTCCAGA	AGTAGTGAGG	2460
25	AGGCTTTTTT	GGAGGCCTAG	GCTTTTGCAA	AAAGCTAGCT	TACAGCTCAG	GGCTGCGATT	2520
	TCGCGCCAAA	CTTGACGGCA	ATCCTAGCGT	GAAGGCTGGT	AGGATTTTAT	CCCCGCTGCC	2580
	ATCATGGTTC	GACCATTGAA	CTGCATCGTC	GCCGTGTCCC	AAAATATGGG	GATTGGCAAG	2640
	AACGGAGACC	TACCCTGGCC	TCCGCTCAGG	AACGAGTTCA	AGTACTTCCA	AAGAATGACC	2700
30	ACAACCTCTT	CAGTGGAAGG	TAAACAGAAT	CTGGTGATTA	TGGGTAGGAA	AACCTGGTTC	2760
	TCCATTCTTG	AGAAGAATCG	ACCTTTAAAG	GACAGAATTA	ATATAGTTCT	CAGTAGAGAA	2820
	CTCAAAGAAC	CACCACGAGG	AGCTCATTTT	CTTGCCAAAA	GTTTGGATGA	TGCCTTAAGA	2880
35	CTTATTGAAC	AACCGGAATT	GGCAAGTAAA	GTAGACATGG	TTTGGATAGT	CGGAGGCAGT	2940
	TCTGTTTACC	AGGAAGCCAT	GAATCAACCA	GGCCACCTCA	GACTCTTTGT	GACAAGGATC	3000
	ATGCAGGAAT	TTGAAAGTGA	CACGTTTTTC	CCAGAAATTG	ATTTGGGGAA	ATATAAACTT	3060
	CTCCCAGAA	ATCCAGGCGT	CCTCTCTGAG	GTCCAGGAGG	AAAAAGGCAT	CAAGTATAAG	3120
40	TTTGAAGTCT	ACGAGAAGAA	AGACTAACAG	GAAGATGCTT	TCAAGTTCTC	TGCTCCCCCTC	3180
	CTAAAGCTAT	GCATTTTTAT	AAGACCATGG	GACTTTTGCT	GGCTTTAGAT	CTTTGTGAAG	3240
	GAACCTTACT	TCTGTGGTGT	GACATAATTG	GACAACTAC	CTACAGAGAT	TTAAAGCTCT	3300
45	AAGGTAAATA	TAAAAATTTT	AAGTGTATAA	TGTGTTAAAC	TACTGATTCT	AATGTTTTGT	3360
	GTATTTTAGA	TTCCAACCTA	TGGAAGTATG	GAATGGGAGC	AGTGGTGGAA	TGCCTTTAAT	3420
	GAGGAAAACC	TGTTTTGCTC	AGAAGAAATG	CCATCTAGTG	ATGATGAGGC	TACTGCTGAC	3480
50	TCTCAACATT	CTACTCCTCC	AAAAAAGAAG	AGAAAGGTAG	AAGACCCCAA	GGACTTTTCT	3540
	TCAGAATTGC	TAAGTTTTTT	GAGTCATGCT	GTGTTTAGTA	ATAGAACTCT	TGCTTGCTTT	3600
	GCTATTTACA	CCACAAAGGA	AAAAGCTGCA	CTGCTATACA	AGAAAATTAT	GGAAAAATAT	3660
55							

EP 0 953 639 A1

	TCTGTAACCT	TTATAAGTAG	GCATAACAGT	TATAATCATA	ACATACTGTT	TTTTCTTACT	3720
	CCACACAGGC	ATAGAGTGTC	TGCTATTAAT	AACTATGCTC	AAAAATTGTG	TACCTTTAGC	3780
5	TTTTTAATTT	GTAAAGGGGT	TAATAAGGAA	TATTTGATGT	ATAGTGCCTT	GACTAGAGAT	3840
	CATAATCAGC	CATACCACAT	TTGTAGAGGT	TTTACTTGCT	TTAAAAAACC	TCCCACACCT	3900
	CCCCCTGAAC	CTGAAACATA	AAATGAATGC	AATTGTTGTT	GTTAACTTGT	TTATTGCAGC	3960
10	TTATAATGGT	TACAAATAAA	GCAATAGCAT	CACAAATTTT	ACAAATAAAG	CATTTTTTTT	4020
	ACTGCATTCT	AGTTGTGGTT	TGTCCAAACT	CATCAATGTA	TCTTATCATG	TCTGGATCTA	4080
	ATAAAAGATA	TTTATTTTCA	TTAGATATGT	GTGTTGGTTT	TTTGTGTGCA	GTGCCTCTAT	4140
	CTGGAGGCCA	GGTAGGGCTG	GCCTTGGGGG	AGGGGGAGGC	CAGAATGACT	CCAAGAGCTA	4200
15	CAGGAAGGCA	GGTCAGAGAC	CCCACTGGAC	AAACAGTGGC	TGGACTCTGC	ACCATAACAC	4260
	ACAATCAACA	GGGGAGTGAG	CTGGAAATTT	GCTAGCGAAT	TCCAGCACAC	TGGCGGCCGT	4320
	TACTAGTTAT	TAATAGTAAT	CAATTACGGG	GTCATTAGTT	CATAGCCCAT	ATATGGAGTT	4380
20	CCGCGTTACA	TAACTTACGG	TAAATGGCCC	GCCTGGCTGA	CCGCCCAACG	ACCCCGCCCC	4440
	ATTGACGTCA	ATAATGACGT	ATGTTCCCAT	AGTAACGCCA	ATAGGGACTT	TCCATTGACG	4500
	TCAATGGGTG	GAGTATTTAC	GGTAAACTGC	CCACTTGGCA	GTACATCAAG	TGTATCATAT	4560
	GCCAAGTACG	CCCCCTATTG	ACGTCAATGA	CGGTAAATGG	CCCGCCTGGC	ATTATGCCCA	4620
25	GTACATGACC	TTATGGGACT	TTCTACTTGT	GCAGTACATC	TACGTATTAG	TCATCGCTAT	4680
	TACCATGGTG	ATGCGGTTTT	GGCAGTACAT	CAATGGGCGT	GGATAGCGGT	TTGACTCACG	4740
	GGGATTTCCA	AGTCTCCACC	CCATTGACGT	CAATGGGAGT	TTGTTTTGGC	ACCAAAATCA	4800
30	ACGGGACTTT	CCAAAATGTC	GTAACAACAT	CGCCCCATTG	ACGCAAATGG	GCGGTAGGCG	4860
	TGTACGGTGG	GAGGTCTATA	TAAGCAGAGC	TCGTTTATGT	AACCGTCAGA	TCGCCTGGAG	4920
	ACGCCATCCA	CGCTGTTTTG	ACCTCCATAG	AAGACACCGG	GACCGATCCA	GCCTCCGCGG	4980
	CCGGGAACGG	TGCATTGGAA	CGCGGATTCC	CCGTGCCAAG	AGTGACGTAA	GTACCGCCTA	5040
35	TAGAGTCTAT	AGGCCACACC	CCTTGGCTTC	TTATGCATGC	TATACTGTTT	TTGGCTTGGG	5100
	GTCTATACAC	CCCCGCTTCC	TCATGTTATA	GGTGATGGTA	TAGCTTAGCC	TATAGGTGTG	5160
	GGTTATTGAC	CATTATTGAC	CACTCCCCTA	TTGGTGACGA	TACTTTCCAT	TACTAATCCA	5220
40	TAACATGGCT	CTTTGCCACA	ACTCTCTTTA	TTGGCTATAT	GCCAATACAC	TGTCCTTCAG	5280
	AGACTGACAC	GGACTCTGTA	TTTTTACAGG	ATGGGGTCTC	ATTTATTATT	TACAAATTCA	5340
	CATATACAAC	ACCACCGTCC	CCAGTGCCCG	CAGTTTTTAT	TAAACATAAC	GTGGGATCTC	5400
	CACGCGAATC	TCGGGTACGT	GTTCCGGACA	TGGGCTCTTC	TCCGGTAGCG	GCGGAGCTTC	5460
45	TACATCCGAG	CCCTGCTCCC	ATGCCTCCAG	CGACTCATGG	TCGCTCGGCA	GCTCCTTGCT	5520
	CCTAACAGTG	GAGGCCAGAC	TTAGGCACAG	CACGATGCCC	ACCACCACCA	GTGTGCCGCA	5580
	CAAGGCCGTG	GCGGTAGGGT	ATGTGTCTGA	AAATGAGCTC	GGGGAGCGGG	CTTGCACCGC	5640
50	TGACGCATTT	GGAAGACTTA	AGGCAGCGGC	AGAAGAAGAT	GCAGGCAGCT	GAGTTGTTGT	5700
	GTTCTGATAA	GAGTCAGAGG	TAACTCCCGT	TGCGGTGCTG	TTAACGGTGG	AGGGCAGTGT	5760
55							

EP 0 953 639 A1

AGTCTGAGCA GTACTCGTTG CTGCCGCGCG CGCCACCAGA CATAATAGCT GACAGACTAA 5820
 CAGACTGTTC CTTTCCATGG GTCTTTTCTG CAGTCACCGT CCTTGACACG CGTCTCGGGA 5880
 5 AGCTTGCCGC CACCATGGGA TGGAGCTGGG TCTTTCTCTT TCTCCTGTCA GGAAGTGCAG 5940
 GTGTCTCTC TGAGGTCCAG CTGCAACAGT CTGGACCTGA GCTGGTGAAG CCTGGGGCTT 6000
 CAGTAAAGAT GTCCTGCAAG ACTTCTAGAT ACACATTAC TGAATACACC ATACACTGGG 6060
 10 TGAGACAGAG CCATGGAAAG AGCCTTGAGT GGATTGGAGG TATTAATCCT AACAAATGGTA 6120
 TTCCTAACTA CAACCAGAAG TTCAAGGGCA GGGCCACATT GACTGTAGGC AAGTCCTCCA 6180
 GCACCGCTA CATGGAGCTC CGCAGCTGA CATCTGAGGA TTCTGCGGTC TATTTCTGTG 6240
 CAAGAAGAAG AATCGCCTAT GGTACGACG AGGGCCATGC TATGGACTAC TGGGGTCAAG 6300
 15 GAACCTCAGT CACCGTCTCC TCAGGTGAGT GGATCCTCTG CGCCTGGGCC CAGCTCTGTC 6360
 CCACACCGCG GTCACATGGC ACCACCTCTC TTGCAGCTC CACCAAGGGC CCATCGGTCT 6420
 TCCCCCTGGC ACCCTCCTCC AAGAGCACCT CTGGGGGCAC AGCGGCCCTG GGCTGCCTGG 6480
 20 TCAAGGACTA CTTCCCCGAA CCGGTGACGG TGTCGTGGAA CTCAGGCGCC CTGACCAGCG 6540
 GCGTGACAC CTTCCCGGCT GTCCTACAGT CCTCAGGACT CTACTCCCTC AGCAGCGTGG 6600
 TGACCGTGCC CTCCAGCAGC TTGGGCACCC AGACCTACAT CTGCAACGTG AATCACAAGC 6660
 25 CCAGCAACAC CAAGGTGGAC AAGAAAGTTG AGCCCAAATC TTGTGACAAA ACTCACACAT 6720
 GCCCACCGTG CCCAGCACCT GAACCTCTGG GGGGACCGTC AGTCTTCTC TTCCCCCAA 6780
 AATCCAAGGA CACCTTCATG ATCTCCCGGA CCCCTGAGGT CACATGCGTG GTGGTGGACG 6840
 TGAGCCACGA AGACCTGAG GTCAAGTTCA ACTGGTACGT GGACGGCGTG GAGGTGCATA 6900
 30 ATGCCAAGAC AAAGCCGCGG GAGGAGCAGT ACAACAGCAC GTACCGGGTG GTCAGCGTCC 6960
 TCACCGTCCT GCACCAGGAC TGGCTGAATG GCAAGGAGTA CAAGTGCAAG GTCTCCAACA 7020
 AAGCCCTCCC AGCCCCATC GAGAAAACCA TCTCCAAAGC CAAAGGGCAG CCCCAGAGAAC 7080
 35 CACAGGTGTA CACCTGCCC CCATCCCGGG AGGAGATGAC CAAGAACCAG GTCAGCCTGA 7140
 CCTGCCTGGT CAAAGGCTTC TATCCAGCG ACATCGCCGT GGAGTGGGAG AGCAATGGGC 7200
 AGCCGGAGAA CAACTACAAG ACCACGCCTC CCGTGCTGGA CTCCGACGGC TCCTTCTTCC 7260
 40 TCTACAGCAA GTCACCGTG GACAAGAGCA GGTGGCAGCA GGGGAACGTC TTCTCATGCT 7320
 CCGTGATGCA TGAGGCTCTG CACAACCACT ACACGCAGAA GAGCCTCTCC CTGTCTCCGG 7380
 GTAAATGAGT GCGACGGCCG GCAAGCCCCG CTCCCCGGG TCTCGCGGTC GCACGAGGAT 7440
 GCTTGGCACG TACCCCTGT ACATACTTCC CGGGCGCCCA GCATGGAAAT AAAGCACCGG 7500
 45 ATCTAATAAA AGATATTTAT TTTCAATTAGA TATGTGTGTT GGTTTTGTGT GTGCAGTGCC 7560
 TCTATCTGGA GGCCAGGTAG GGCTGGCCTT GGGGGAGGGG GAGGCCAGAA TGAATCCAAG 7620
 AGCTACAGGA AGGCAGGTCA GAGACCCAC TGGACAAACA GTGGCTGGAC TCTGCACCAT 7680
 50 AACACACAAT CAACAGGGGA GTGAGCTGGA AATTTGCTAG CGAATTAATT C 7731

(2) INFORMATION FOR SEQ ID NO: 30:

(i) SEQUENCE CHARACTERISTICS:

55

(A) LENGTH: 472 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 30:

Met	Gly	Trp	Ser	Trp	Val	Phe	Leu	Phe	Leu	Leu	Ser	Gly	Thr	Ala	Gly	1	5	10	15
Val	Leu	Ser	Glu	Val	Gln	Leu	Gln	Gln	Ser	Gly	Pro	Glu	Leu	Val	Lys	20	25	30	
Pro	Gly	Ala	Ser	Val	Lys	Met	Ser	Cys	Lys	Thr	Ser	Arg	Tyr	Thr	Phe	35	40	45	
Thr	Glu	Tyr	Thr	Ile	His	Trp	Val	Arg	Gln	Ser	His	Gly	Lys	Ser	Leu	50	55	60	
Glu	Trp	Ile	Gly	Gly	Ile	Asn	Pro	Asn	Asn	Gly	Ile	Pro	Asn	Tyr	Asn	65	70	75	80
Gln	Lys	Phe	Lys	Gly	Arg	Ala	Thr	Leu	Thr	Val	Gly	Lys	Ser	Ser	Ser	85	90	95	
Thr	Ala	Tyr	Met	Glu	Leu	Arg	Ser	Leu	Thr	Ser	Glu	Asp	Ser	Ala	Val	100	105	110	
Tyr	Phe	Cys	Ala	Arg	Arg	Arg	Ile	Ala	Tyr	Gly	Tyr	Asp	Glu	Gly	His	115	120	125	
Ala	Met	Asp	Tyr	Trp	Gly	Gln	Gly	Thr	Ser	Val	Thr	Val	Ser	Ser	Ser	130	135	140	
Thr	Lys	Gly	Pro	Ser	Val	Phe	Pro	Leu	Ala	Pro	Ser	Ser	Lys	Ser	Thr	145	150	155	160
Ser	Gly	Gly	Thr	Ala	Ala	Leu	Gly	Cys	Leu	Val	Lys	Asp	Tyr	Phe	Pro	165	170	175	
Glu	Pro	Val	Thr	Val	Ser	Trp	Asn	Ser	Gly	Ala	Leu	Thr	Ser	Gly	Val	180	185	190	
His	Thr	Phe	Pro	Ala	Val	Leu	Gln	Ser	Ser	Gly	Leu	Tyr	Ser	Leu	Ser	195	200	205	
Ser	Val	Val	Thr	Val	Pro	Ser	Ser	Ser	Leu	Gly	Thr	Gln	Thr	Tyr	Ile	210	215	220	
Cys	Asn	Val	Asn	His	Lys	Pro	Ser	Asn	Thr	Lys	Val	Asp	Lys	Lys	Val	225	230	235	240
Glu	Pro	Lys	Ser	Cys	Asp	Lys	Thr	His	Thr	Cys	Pro	Pro	Cys	Pro	Ala	245	250	255	
Pro	Glu	Leu	Leu	Gly	Gly	Pro	Ser	Val	Phe	Leu	Phe	Pro	Pro	Lys	Pro	260	265	270	
Lys	Asp	Thr	Leu	Met	Ile	Ser	Arg	Thr	Pro	Glu	Val	Thr	Cys	Val	Val	275	280	285	
Val	Asp	Val	Ser	His	Glu	Asp	Pro	Glu	Val	Lys	Phe	Asn	Trp	Tyr	Val	290	295	300	
Asp	Gly	Val	Glu	Val	His	Asn	Ala	Lys	Thr	Lys	Pro	Arg	Glu	Glu	Gln				

EP 0 953 639 A1

		305				310				315				320			
5		Tyr	Asn	Ser	Thr	Tyr 325	Arg	Val	Val	Ser	Val 330	Leu	Thr	Val	Leu	His 335	Gln
		Asp	Trp	Leu	Asn 340	Gly	Lys	Glu	Tyr	Lys 345	Cys	Lys	Val	Ser	Asn 350	Lys	Ala
		Leu	Pro	Ala 355	Pro	Ile	Glu	Lys	Thr 360	Ile	Ser	Lys	Ala 365	Lys	Gly	Gln	Pro
10		Arg	Glu 370	Pro	Gln	Val	Tyr	Thr 375	Leu	Pro	Pro	Ser	Arg 380	Glu	Glu	Met	Thr
		Lys 385	Asn	Gln	Val	Ser	Leu 390	Thr	Cys	Leu	Val	Lys 395	Gly	Phe	Tyr	Pro	Ser 400
15		Asp	Ile	Ala	Val	Glu 405	Trp	Glu	Ser	Asn	Gly 410	Gln	Pro	Glu	Asn 415	Asn	Tyr
		Lys	Thr	Thr	Pro 420	Pro	Val	Leu	Asp	Ser 425	Asp	Gly	Ser	Phe 430	Phe	Leu	Tyr
20		Ser	Lys	Leu 435	Thr	Val	Asp	Lys	Ser 440	Arg	Trp	Gln	Gln	Gly 445	Asn	Val	Phe
		Ser 450	Cys	Ser	Val	Met	His	Glu 455	Ala	Leu	His	Asn 460	His	Tyr	Thr	Gln	Lys
		Ser 465	Leu	Ser	Leu	Ser	Pro 470	Gly	Lys								

(2) INFORMATION FOR SEQ ID NO: 31:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 339 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 31:

GACATTGTGA	TGACCCAATC	TCCAGACTCT	TTGGCTGTGT	CTCTAGGGGA	GAGGGCCACC	60
ATCAACTGCA	AGTCCAGTCA	GAGCCTTTTA	TATTCTAGAA	ATCAAAAGAA	CTACTTGGCC	120
TGGTATCAGC	AGAAACCAGG	ACAGCCACCC	AAACTCCTCA	TCTTTTGGGC	TAGCACTAGG	180
GAATCTGGGG	TACCTGATAG	GTTCA GTGGC	AGTGGGTTTG	GGACAGACTT	CACCCTCACC	240
ATTAGCAGCC	TGCAGGCTGA	AGATGTGGCA	GTTTATTACT	GTCAGCAATA	TTTTAGCTAT	300
CCGCTCACGT	TCGGACAAGG	GACCAAGGTG	GAAATAAAAA			339

(2) INFORMATION FOR SEQ ID NO: 32:

(i) SEQUENCE CHARACTERISTICS:

- SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 113 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 32:

Asp Ile Val Met Thr Gln Ser Pro Asp Ser Leu Ala Val Ser Leu Gly
1 5 10 15
Glu Arg Ala Thr Ile Asn Cys Lys Ser Ser Gln Ser Leu Leu Tyr Ser
20 25 30
Arg Asn Gln Lys Asn Tyr Leu Ala Trp Tyr Gln Gln Lys Pro Gly Gln
35 40 45
Pro Pro Lys Leu Leu Ile Phe Trp Ala Ser Thr Arg Glu Ser Gly Val
50 55 60
Pro Asp Arg Phe Ser Gly Ser Gly Phe Gly Thr Asp Phe Thr Leu Thr
65 70 75 80
Ile Ser Ser Leu Gln Ala Glu Asp Val Ala Val Tyr Tyr Cys Gln Gln
85 90 95
Tyr Phe Ser Tyr Pro Leu Thr Phe Gly Gln Gly Thr Lys Val Glu Ile
100 105 110
Lys

(2) INFORMATION FOR SEQ ID NO: 33:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 113 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 33:

Asp Ile Val Met Thr Gln Ser Pro Asp Ser Leu Ala Val Ser Leu Gly
1 5 10 15
Glu Arg Ala Thr Ile Asn Cys Lys Ser Ser Gln Ser Leu Leu Tyr Ser
20 25 30
Arg Asn Gln Lys Asn Tyr Leu Ala Trp Phe Gln Gln Lys Pro Gly Gln
35 40 45
Pro Pro Lys Leu Leu Ile Phe Trp Ala Ser Thr Arg Glu Ser Gly Val
50 55 60
Pro Asp Arg Phe Ser Gly Ser Gly Phe Gly Thr Asp Phe Thr Leu Thr
65 70 75 80
Ile Ser Ser Leu Gln Ala Glu Asp Val Ala Val Tyr Asp Cys Gln Gln
85 90 95
Tyr Phe Ser Tyr Pro Leu Thr Phe Gly Gln Gly Thr Lys Val Glu Ile
100 105 110
Lys

(2) INFORMATION FOR SEQ ID NO: 34:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 113 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 34:

Asp	Ile	Val	Met	Thr	Gln	Ser	Pro	Asp	Ser	Leu	Ala	Val	Ser	Leu	Gly
1				5				10						15	
Glu	Arg	Ala	Thr	Ile	Asn	Cys	Lys	Ser	Gln	Ser	Leu	Leu	Tyr	Ser	
			20					25				30			
Arg	Asn	Gln	Lys	Asn	Tyr	Leu	Ala	Trp	Tyr	Gln	Gln	Lys	Pro	Gly	Gln
		35					40					45			
Pro	Pro	Lys	Leu	Leu	Ile	Tyr	Trp	Ala	Ser	Thr	Arg	Glu	Ser	Gly	Val
		50				55					60				
Pro	Asp	Arg	Phe	Ser	Gly	Ser	Gly	Phe	Gly	Thr	Asp	Phe	Thr	Leu	Thr
					70					75				80	
Ile	Ser	Ser	Leu	Gln	Ala	Glu	Asp	Val	Ala	Val	Tyr	Tyr	Cys	Gln	Gln
				85					90					95	
Tyr	Phe	Ser	Tyr	Pro	Leu	Thr	Phe	Gly	Gln	Gly	Thr	Lys	Val	Glu	Ile
			100					105					110		

Lys

(2) INFORMATION FOR SEQ ID NO: 35:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 8068 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 35:

GAATTCCAGC	ACACTGGCGG	CCGTTACTAG	TTATTAATAG	TAATCAATTA	CGGGGTCATT	60
AGTTCATAGC	CCATATATGG	AGTTCGCGT	TACATAACTT	ACGGTAAATG	GCCCCGCTGG	120
CTGACCGCCC	AACGACCCCC	GCCCATTGAC	GTCAATAATG	ACGTATGTTT	CCATAGTAAC	180
GCCAATAGGG	ACTTTCCATT	GACGTCAATG	GGTGGAGTAT	TTACGGTAAA	CTGCCCCACTT	240
GGCAGTACAT	CAAGTGTATC	ATATGCCAAG	TACGCCCCCT	ATTGACGTCA	ATGACGGTAA	300
ATGGCCCGCC	TGGCATTATG	CCCAGTACAT	GACCTTATGG	GACTTTCCTA	CTTGGCAGTA	360
CATCTACGTA	TTAGTCATCG	CTATTACCAT	GGTGATGCGG	TTTTGGCAGT	ACATCAATGG	420
GCGTGGATAG	CGGTTTGACT	CACGGGGATT	TCCAAGTCTC	CACCCCATTG	ACGTCAATGG	480
GAGTTTGTGT	TGGCACCAAA	ATCAACGGGA	CTTTCCAAAA	TGTCGTAAAC	ACTCCGCCCC	540
ATTGACGCAA	ATGGGCGGTA	GGCGTGTACG	GTGGGAGGTC	TATATAAGCA	GAGCTCGTTT	600

EP 0 953 639 A1

	AGTGAACCGT	CAGATCGCCT	GGAGACGCCA	TCCACGCTGT	TTTGACCTCC	ATAGAAGACA	660
	CCGGGACCGA	TCCAGCCTCC	GCGGCCGGGA	ACGGTGCATT	GGAACGCGGA	TTCCCCGTGC	720
5	CAAGAGTGAC	GTAAGTACCG	CCTATAGAGT	CTATAGGCC	ACCCCTTGG	CTTCTTATGC	780
	ATGCTATACT	GTTTTTGGCT	TGGGGTCTAT	ACACCCCCGC	TTCTCATGT	TATAGGTGAT	840
	GGTATAGCTT	AGCCTATAGG	TGTGGGTAT	TGACCATAT	TGACCACTCC	CCTATTGGTG	900
10	ACGATACTTT	CCATTACTAA	TCCATAACAT	GGCTCTTTGC	CACAACTCTC	TTTATTGGCT	960
	ATATGCCAAT	ACACTGTCCT	TCAGAGACTG	ACACGGACTC	TGTATTTTTA	CAGGATGGGG	1020
	TCTCATTTAT	TATTTACAAA	TTACATATA	CAACACCACC	GTCCCCAGTG	CCCGCAGTTT	1080
	TTATTAAACA	TAACGTGGGA	TCTCCACGCG	AATCTCGGGT	ACGTGTTCCG	GACATGGGCT	1140
15	CTTCTCCGGT	AGCGGCGGAG	CTTCTACATC	CGAGCCCTGC	TCCCATGCCT	CCAGCGACTC	1200
	ATGGTCGCTC	GGCAGCTCCT	TGCTCCTAAC	AGTGGAGGCC	AGACTTAGGC	ACAGCACGAT	1260
	GCCCACCACC	ACCACTGTGC	CGCACAAAGC	CGTGGCGGTA	GGGTATGTGT	CTGAAAATGA	1320
20	GCTCGGGGAG	CGGGCTTGCA	CCGCTGACGC	ATTTGGAAGA	CTTAAGGCAG	CGGCAGAAGA	1380
	AGATGCAGGC	AGCTGAGTTG	TTGTGTTCTG	ATAAGAGTCA	GAGGTAAGTC	CCGTTGCGGT	1440
	GCTGTTAACG	GTGGAGGGCA	GTGTAGTCTG	AGCAGTACTC	GTTGCTGCCG	CGCGCGCCAC	1500
	CAGACATAAT	AGCTGACAGA	CTAACAGACT	GTTCCCTTTC	ATGGGTCTTT	TCTGCAGTCA	1560
25	CCGTCCCTGA	CACGCGTCTC	GGGAAGCTTG	CCGCCACCAT	GGAGACAGAC	ACACTCCTGC	1620
	TATGGGTGCT	GCTGCTCTGG	GTTCCAGGTT	CCTCCGAGGA	CATTGTGATG	ACCCAATCTC	1680
	CAGACTCTTT	GGCTGTGTCT	CTAGGGGAGA	GGGCCACCAT	CAACTGCAAG	TCCAGTCAGA	1740
30	GCCTTTTATA	TTCTAGAAAT	CAAAAGAACT	ACTTGGCCTG	GTATCAGCAG	AAACCAGGAC	1800
	AGCCACCCAA	ACTCCTCATC	TTTTGGGCTA	GCACTAGGGA	ATCTGGGGTA	CCTGATAGGT	1860
	TCAGTGGCAG	TGGGTTTGGG	ACAGACTTCA	CCCTCACCAT	TAGCAGCCTG	CAGGCTGAAG	1920
35	ATGTGGCAGT	TTATTACTGT	CAGCAATATT	TTAGCTATCC	GCTCACGTTT	GGACAAGGGA	1980
	CCAAGGTGGA	AATAAAACGT	GAGTGGATCC	ATCTGGGATA	AGCATGCTGT	TTTCTGTCTG	2040
	TCCCTAACAT	GCCCTGTGAT	TATGCGCAA	CAACACACCC	AAGGGCAGAA	CTTTGTTACT	2100
	TAAACACCAT	CCTGTTTGCT	TCTTTCCTCA	GGAACGTGG	CTGCACCATC	TGTCTTCATC	2160
40	TTCCCGCCAT	CTGATGAGCA	GTTGAAATCT	GGAAGTGCCT	CTGTTGTGTG	CCTGCTGAAT	2220
	AACTTCTATC	CCAGAGAGGC	CAAAGTACAG	TGGAAGGTGG	ATAACGCCCT	CCAATCGGGT	2280
	AACTCCAGG	AGAGTGTAC	AGAGCAGGAC	AGCAAGGACA	GCACCTACAG	CCTCAGCAGC	2340
45	ACCCTGACGC	TGAGCAAAGC	AGACTACGAG	AAACACAAAG	TCTACGCCTG	CGAAGTCACC	2400
	CATCAGGGCC	TGAGCTCGCC	CGTCACAAAG	AGCTTCAACA	GGGGAGAGTG	TTAGAGGGAG	2460
	AAGTGCCCCC	ACCTGCTCCT	CAGTTCCAGC	CTGACCCCT	CCCATCCTTT	GGCCTCTGAC	2520
	CCTTTTTCCA	CAGGGGACCT	ACCCCTATTG	CGGTCCTCCA	GCTCATCTTT	CACCTCACCC	2580
50	CCCTCCTCCT	CCTTGGCTTT	AATTATGCTA	ATGTTGGAGG	AGAATGAATA	AATAAAGTGA	2640
	ATCTTTGCAC	CTGTGGTGGA	TCTAATAAAA	GATATTTATT	TTCATTAGAT	ATGTGTGTTG	2700

55

EP 0 953 639 A1

	GTTTTTTGTG	TGCAGTGCCT	CTATCTGGAG	GCCAGGTAGG	GCTGGCCTTG	GGGGAGGGGG	2760
	AGGCCAGAAT	GACTCCAAGA	GCTACAGGAA	GGCAGGTCAG	AGACCCCACT	GGACAAACAG	2820
5	TGGCTGGACT	CTGCACCATA	ACACACAATC	AACAGGGGAG	TGAGCTGGAA	ATTTGCTAGC	2880
	GAATTCCTGA	AGACGAAAGG	GCCTCGTGAT	ACGCCTATTT	TTATAGGTTA	ATGTCATGAT	2940
	AATAATGGTT	TCTTAGACGT	CAGGTGGCAC	TTTTCGGGGA	AATGTGCGCG	GAACCCCTAT	3000
10	TTGTTTATTT	TTCTAAATAC	ATTCAAATAT	GTATCCGCTC	ATGAGACAAT	AACCCTGATA	3060
	AATGCTTCAA	TAATATTGAA	AAAGGAAGAG	TATGAGTATT	CAACATTTCC	GTGTCGCCCT	3120
	TATTCCTTTT	TTTGCGGCAT	TTTGCCTTCC	TGTTTTTGCT	CACCCAGAAA	CGCTGGTGAA	3180
	AGTAAAAGAT	GCTGAAGATC	AGTTGGGTGC	ACGAGTGGGT	TACATCGAAC	TGGATCTCAA	3240
15	CAGCGGTAAG	ATCCTTGAGA	GTTTTCGCCC	CGAAGAACGT	TTTCCAATGA	TGAGCACTTT	3300
	TAAAGTTCTG	CTATGTGGCG	CGGTATTATC	CCGTGTTGAC	GCCGGGCAAG	AGCAACTCGG	3360
	TCGCCGCATA	CACTATTCTC	AGAATGACTT	GGTTGAGTAC	TCACCAGTCA	CAGAAAAGCA	3420
20	TCTTACGGAT	GGCATGACAG	TAAGAGAATT	ATGCAGTGCT	GCCATAACCA	TGAGTGATAA	3480
	CACTGCGGCC	AACTTACTTC	TGACAACGAT	CGGAGGACCG	AAGGAGCTAA	CCGCTTTTTT	3540
	GCACAACATG	GGGGATCATG	TAACTCGCCT	TGATCGTTGG	GAACCGGAGC	TGAATGAAGC	3600
25	CATACCAAAC	GACGAGCGTG	ACACCACGAT	GCCTGCAGCA	ATGGCAACAA	CGTTGCGCAA	3660
	ACTATTAACT	GGCGAACTAC	TTACTCTAGC	TTCCCGGCAA	CAATTAATAG	ACTGGATGGA	3720
	GGCGGATAAA	GTTGCAGGAC	CACTTCTGCG	CTCGGCCCTT	CCGGCTGGCT	GGTTTATTGC	3780
	TGATAAATCT	GGAGCCGGTG	AGCGTGGGTC	TCGCGGTATC	ATTGCAGCAC	TGGGGCCAGA	3840
30	TGGTAAGCCC	TCCCGTATCG	TAGTTATCTA	CACGACGGGG	AGTCAGGCAA	CTATGGATGA	3900
	ACGAAATAGA	CAGATCGCTG	AGATAGGTGC	CTCACTGATT	AAGCATTGGT	AACTGTCAGA	3960
	CCAAGTTTAC	TCATATATAC	TTTAGATTGA	TTTAAACTTT	CATTTTAAAT	TTAAAGGAT	4020
35	CTAGGTGAAG	ATCCTTTTTC	ATAATCTCAT	GACCAAAATC	CCTTAACGTG	AGTTTTCGTT	4080
	CCACTGAGCG	TCAGACCCCG	TAGAAAAGAT	CAAAGGATCT	TCTTGAGATC	CTTTTTTTCT	4140
	GCGCGTAATC	TGCTGCTTGC	AAACAAAAAA	ACCACCCTA	CCAGCGGTGG	TTTGTTTGCC	4200
	GGATCAAGAG	CTACCAACTC	TTTTTCCGAA	GGTAACTGGC	TTCAGCAGAG	CGCAGATACC	4260
40	AAATACTGTC	CTTCTAGTGT	AGCCGTAGTT	AGGCCACCAC	TTCAAGAACT	CTGTAGCACC	4320
	GCCTACATAC	CTCGCTCTGC	TAATCCTGTT	ACCAGTGGCT	GCTGCCAGTG	GCGATAAGTC	4380
	GTGTCTTACC	GGGTTGGACT	CAAGACGATA	GTTACCGGAT	AAGGCGCAGC	GGTCGGGCTG	4440
45	AACGGGGGGT	TCGTGCACAC	AGCCCAGCTT	GGAGCGAACG	ACCTACACCG	AACTGAGATA	4500
	CCTACAGCGT	GAGCTATGAG	AAAGCGCCAC	GCTTCCCGAA	GGGAGAAAGG	CGGACAGGTA	4560
	TCCGGTAAGC	GGCAGGGTCG	GAACAGGAGA	GCGCACGAGG	GAGCTTCCAG	GGGGAAACGC	4620
	CTGGTATCTT	TATAGTCCTG	TCGGGTTTCG	CCACCTCTGA	CTTGAGCGTC	GATTTTTGTG	4680
50	ATGCTCGTCA	GGGGGGCGGA	GCCTATGGAA	AAACGCCAGC	AACGCGGCCT	TTTTACGGTT	4740
	CCTGGCCTTT	TGCTGGCCTT	TTGCTCACAT	GTTCTTTTCT	GCGTTATCCC	CTGATTCTGT	4800
55							

EP 0 953 639 A1

	GGATAACCGT ATTACCGCCT TTGAGTGAGC TGATACCGCT CGCCGCAGCC GAACGACCGA	4860
	GCGCAGCGAG TCAGTGAGCG AGGAAGCGGA AGAGCGCCTG ATGCGGTATT TTCTCCTTAC	4920
5	GCATCTGTGC GGTATTTTAC ACCGCATATG GTGCACTCTC AGTACAATCT GCTCTGATGC	4980
	CGCATAGTTA AGCCAGTATA CACTCCGCTA TCGCTACGTG ACTGGGTCAT GGCTGCGCCC	5040
	CGACACCCGC CAACACCCGC TGACGCGCCC TGACGGGCTT GTCTGCTCCC GGCATCCGCT	5100
10	TACAGACAAG CTGTGACCGT CTCGGGAGC TGCATGTGTC AGAGGTTTTT ACCGTCATCA	5160
	CCGAAACGCG CGAGGCAGCT GTGGAATGTG TGTCAGTTAG GGTGTGAAA GTCCCCAGGC	5220
	TCCCCAGCAG GCAGAAGTAT GCAAAGCATG CATCTCAATT AGTCAGCAAC CAGGCTCCCC	5280
	AGCAGGCAGA AGTATGCAAA GCATGCATCT CAATTAGTCA GCAACCATAG TCCCGCCCCT	5340
15	AACTCCGCCC ATCCCGCCCC TAACTCCGCC CAGTTCCGCC CATTCTCCGC CCCATGGCTG	5400
	ACTAATTTTT TTTATTTATG CAGAGGCCGA GGCCGCCCTCG GCCTCTGAGC TATTCCAGAA	5460
	GTAGTGAGGA GGCTTTTGTG GAGGCCTAGG CTTTGTGAAA AAGCTAGCTT CACGCTGCCG	5520
20	CAAGCACTCA GGGCGCAAGG GCTGCTAAAG GAAGCGGAAC ACGTAGAAAG CCAGTCCGCA	5580
	GAAACGGTGC TGACCCCGGA TGAATGTCAG CTACTGGGCT ATCTGGACAA GGGAAAACGC	5640
	AAGCGCAAAG AGAAAGCAGG TAGCTTGCAG TGGGCTTACA TGGCGATAGC TAGACTGGGC	5700
	GGTTTATATG ACAGCAAGCG AACCGBAATT GCCAGCTGGG GCGCCCTCTG GTAAGGTTGG	5760
25	GAAGCCCTGC AAAGTAAACT GGATGGCTTT CTTGCCGCCA AGGATCTGAT GGCAGGGGG	5820
	ATCAAGATCT GATCAAGAGA CAGGATGAGG ATCGTTTTCG ATGATTGAAC AAGATGGATT	5880
	GCACGCAGGT TCTCCGGCCG CTTGGGTGGA GAGGCTATTG GGCTATGACT GGGACAACA	5940
30	GACAATCGGC TGCTCTGATG CCGCCGTGTT CCGGCTGTCA GCGCAGGGGC GCCCGGTTCT	6000
	TTTTGTCAAG ACCGACCTGT CCGGTGCCCT GAATGAACTG CAGGACGAGG CAGCGCGGCT	6060
	ATCGTGGCTG GCCACGACGG GCGTTCCTTG CCGAGCTGTG CTCGACGTTG TCACTGAAGC	6120
35	GGGAAGGGAC TGGCTGCTAT TGGGCGAAGT GCCGGGGCAG GATCTCCTGT CATCTCACCT	6180
	TGCTCCTGCC GAGAAAGTAT CCATCATGGC TGATGCAATG CGGCGGCTGC ATACGCTTGA	6240
	TCCGGCTACC TGCCCATTCG ACCACCAAGC GAAACATCGC ATCGAGCGAG CACGTACTCG	6300
	GATGGAAGCC GGTCTTGTG ATCAGGATGA TCTGGACGAA GAGCATCAGG GGCTCGCGCC	6360
40	AGCCGAACTG TTCGCCAGGC TCAAGGCGCG CATGCCCGAC GGCAGAGGATC TCGTCGTGAC	6420
	CCATGGCGAT GCCTGCTTGC CGAATATCAT GGTGGAAAAT GGCCGCTTTT CTGGATTTCAT	6480
	CGACTGTGGC CGGCTGGGTG TGGCGGACCG CTATCAGGAC ATAGCGTTGG CTACCCGTGA	6540
45	TAITGCTGAA GAGCTTGGCG GCGAATGGGC TGACCGCTTC CTCGTGCTTT ACGGTATCGC	6600
	CGCTCCCGAT TCGCAGCGCA TCGCCTTCTA TCGCCTTCTT GACGAGTTCT TCTGAGCGGG	6660
	ACTCTGGGGT TCGAAATGAC CGACCAAGCG ACGCCCAACC TGCCATCACG AGATTTTCGAT	6720
50	TCCACCGCCG CCTTCTATGA AAGGTTGGGC TTCGGAATCG TTTTCCGGGA CGCCGGCTGG	6780
	ATGATCCTCC AGCGCGGGGA TCTCATGCTG GAGTTCTTCG CCCACCCCGG GCTCGATCCC	6840
	CTCGCGAGTT GGTTCAGCTG CTGCCGTGAGG CTGGACGACC TCGCGGAGTT CTACCGGCAG	6900
55		

5 TGCAAATCCG TCGGCATCCA GGAAACCAGC AGCGGCTATC CGCGCATCCA TGCCCCCGAA 6960
 CTGCAGGAGT GGGGAGGCAC GATGGCCGCT TTGGTCCCGG ATCTTTGTGA AGGAACCTTA 7020
 CTTCTGTGGT GTGACATAAT TGGACAACT ACCTACAGAG ATTTAAAGCT CTAAGGTAAA 7080
 TATAAAATTT TTAAGTGTAT AATGTGTAA ACTACTGATT CTAATTGTTT GTGTATTTTA 7140
 GATTCCAACC TATGGAAC TGAAATGGGA GCAGTGGTGG AATGCCTTTA ATGAGGAAAA 7200
 10 CCTGTTTTGC TCAGAAGAAA TGCCATCTAG TGATGATGAG GCTACTGCTG ACTCTCAACA 7260
 TTCTACTCCT CCAAAAAGA AGAGAAAGGT AGAAGACCCC AAGGACTTTC CTTCAGAATT 7320
 GCTAAGTTTT TTGAGTCATG CTGTGTTTAG TAATAGAACT CTTGCTTGCT TTGCTATTTA 7380
 CACCACAAAG GAAAAAGCTG CACTGCTATA CAAGAAAATT ATGGAAAAAT ATTCTGTAAC 7440
 15 CTTTATAAGT AGGCATAACA GTTATAATCA TAACATACTG TTTTCTCTTA CTCCACACAG 7500
 GCATAGAGTG TCTGCTATTA ATAACATATG TCAAAAATTG TGTACCTTTA GCTTTTAAAT 7560
 TTGTAAAGGG GTTAATAAGG AATATTTGAT GTATAGTGCC TTGACTAGAG ATCATAATCA 7620
 20 GCCATACCAC ATTTGTAGAG GTTTTACTTG CTTTAAAAAA CCTCCACAC CTCCCCCTGA 7680
 ACCTGAAACA TAAATGAAT GCAATTGTTG TTGTTAACTT GTTTATTGCA GCTTATAATG 7740
 GTTACAAATA AAGCAATAGC ATCACAAATT TCACAAATAA AGCATTTTTT TCACTGCATT 7800
 25 CTAGTTGTGG TTTGTCCAAA CTCATCAATG TATCTTATCA TGTCTGGATC TAATAAAGA 7860
 TATTTATTTT CATTAGATAT GTGTGTTGGT TTTTGTGTG CAGTGCCTCT ATCTGGAGGC 7920
 CAGGTAGGGC TGGCCTTGGG GGAGGGGGAG GCCAGAATGA CTCCAAGAGC TACAGGAAGG 7980
 CAGGTCAGAG ACCCCACTGG ACAAACAGTG GCTGGACTCT GCACCATAAC ACACAATCAA 8040
 30 CAGGGGAGTG AGCTGGAAT TTGCTAGC 8068

(2) INFORMATION FOR SEQ ID NO: 36:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 234 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 36:

Met Glu Thr Asp Thr Leu Leu Leu Trp Val Leu Leu Leu Trp Val Pro
 1 5 10 15
 45 Gly Ser Ser Gly Asp Ile Val Met Thr Gln Ser Pro Asp Ser Leu Ala
 20 25 30
 Val Ser Leu Gly Glu Arg Ala Thr Ile Asn Cys Lys Ser Ser Gln Ser
 35 40 45
 50 Leu Leu Tyr Ser Arg Asn Gln Lys Asn Tyr Leu Ala Trp Tyr Gln Gln
 50 55 60
 Lys Pro Gly Gln Pro Pro Lys Leu Leu Ile Phe Trp Ala Ser Thr Arg
 65 70 75 80

5		Glu	Ser	Gly	Val	Pro 85	Asp	Arg	Phe	Ser	Gly 90	Ser	Gly	Phe	Gly	Thr 95	Asp
		Phe	Thr	Leu	Thr 100	Ile	Ser	Ser	Leu	Gln 105	Ala	Glu	Asp	Val	Ala 110	Val	Tyr
		Tyr	Cys	Gln 115	Gln	Tyr	Phe	Ser	Tyr 120	Pro	Leu	Thr	Phe	Gly 125	Gln	Gly	Thr
10		Lys	Val 130	Glu	Ile	Lys	Arg	Val 135	Phe	Ile	Phe	Pro	Pro 140	Ser	Asp	Glu	Gln
		Leu 145	Lys	Ser	Gly	Thr	Ala 150	Ser	Val	Val	Cys 155	Leu	Leu	Asn	Asn	Phe	Tyr 160
15		Pro	Arg	Glu	Ala	Lys 165	Val	Gln	Trp	Lys	Val 170	Asp	Asn	Ala	Leu	Gln 175	Ser
		Gly	Asn	Ser	Gln 180	Glu	Ser	Val	Thr	Glu 185	Gln	Asp	Ser	Lys	Asp 190	Ser	Thr
		Tyr	Ser	Leu 195	Ser	Ser	Thr	Leu	Thr 200	Leu	Ser	Lys	Ala	Asp 205	Tyr	Glu	Lys
20		His	Lys 210	Val	Tyr	Ala	Cys	Glu 215	Val	Thr	His	Gln	Gly 220	Leu	Ser	Ser	Pro
		Val 225	Thr	Lys	Ser	Phe	Asn 230	Arg	Gly	Glu	Cys						

(2) INFORMATION FOR SEQ ID NO: 37:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 372 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 37:

CAGGTGCAAC	TAGTGCAGTC	CGGCGCCGAA	GTGAAGAAAC	CCGGTGCTTC	CGTGAAAGTC	60
AGCTGTAAAA	CTAGTAGATA	CACCTTCACT	GAATACACCA	TACACTGGGT	TAGACAGGCC	120
CCTGGCCAAA	GGCTGGAGTG	GATAGGAGGT	ATTAATCCTA	ACAATGGTAT	TCCTAACTAC	180
AACCAGAAGT	TCAAGGGCCG	GGCCACCTTG	ACCGTAGGCA	AGTCTGCCAG	CACCGCCTAC	240
ATGGAAGTGT	CCAGCCTGCG	CTCCGAGGAC	ACTGCAGTCT	ACTACTGCGC	CAGAAGAAGA	300
ATCGCCTATG	GTTACGACGA	GGGCCATGCT	ATGGACTACT	GGGGTCAAGG	AACCCTTGTC	360
ACCGTCTCCT	CA					372

(2) INFORMATION FOR SEQ ID NO: 38:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 124 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 38:

5 Gln Val Gln Leu Val Gln Ser Gly Ala Glu Val Lys Lys Pro Gly Ala
 1 5 10 15
 Ser Val Lys Val Ser Cys Lys Thr Ser Arg Tyr Thr Phe Thr Glu Tyr
 20 25 30
 Thr Ile His Trp Val Arg Gln Ala Pro Gly Gln Arg Leu Glu Trp Ile
 35 40 45
 10 Gly Gly Ile Asn Pro Asn Asn Gly Ile Pro Asn Tyr Asn Gln Lys Phe
 50 55 60
 Lys Gly Arg Ala Thr Leu Thr Val Gly Lys Ser Ala Ser Thr Ala Tyr
 65 70 75 80
 15 Met Glu Leu Ser Ser Leu Arg Ser Glu Asp Thr Ala Val Tyr Tyr Cys
 85 90 95
 Ala Arg Arg Arg Ile Ala Tyr Gly Tyr Asp Glu Gly His Ala Met Asp
 100 105 110
 20 Tyr Trp Gly Gln Gly Thr Leu Val Thr Val Ser Ser
 115 120

(2) INFORMATION FOR SEQ ID NO: 39:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 124 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 39:

35 Gln Val Gln Leu Val Gln Ser Gly Ala Glu Val Lys Lys Pro Gly Ala
 1 5 10 15
 Ser Val Lys Val Ser Cys Lys Thr Ser Arg Tyr Thr Phe Thr Glu Tyr
 20 25 30
 Thr Ile His Trp Val Arg Gln Ala Pro Gly Gln Arg Leu Glu Trp Ile
 35 40 45
 40 Gly Gly Ile Asn Pro Asn Asn Gly Ile Pro Asn Tyr Asn Gln Lys Phe
 50 55 60
 Lys Gly Arg Ala Thr Leu Thr Val Gly Lys Ser Ala Ser Thr Ala Tyr
 65 70 75 80
 45 Met Glu Leu Ser Ser Leu Arg Ser Glu Asp Thr Ala Val Tyr Phe Cys
 85 90 95
 Ala Arg Arg Arg Ile Ala Tyr Gly Tyr Asp Glu Gly His Ala Met Asp
 100 105 110
 50 Tyr Trp Gly Gln Gly Thr Leu Val Thr Val Ser Ser
 115 120

(2) INFORMATION FOR SEQ ID NO: 40:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 124 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 40:

Gln	Val	Gln	Leu	Val	Gln	Ser	Gly	Ala	Glu	Val	Lys	Lys	Pro	Gly	Ala	1	5	10	15
Ser	Val	Lys	Val	Ser	Cys	Lys	Thr	Ser	Arg	Tyr	Thr	Phe	Thr	Glu	Tyr	20	25	30	
Thr	Ile	His	Trp	Val	Arg	Gln	Ala	Pro	Gly	Gln	Arg	Leu	Glu	Trp	Ile	35	40	45	
Gly	Gly	Ile	Asn	Pro	Asn	Asn	Gly	Ile	Pro	Asn	Tyr	Asn	Gln	Lys	Phe	50	55	60	
Lys	Gly	Arg	Val	Thr	Ile	Thr	Val	Asp	Thr	Ser	Ala	Ser	Thr	Ala	Tyr	65	70	75	80
Met	Glu	Leu	Ser	Ser	Leu	Arg	Ser	Glu	Asp	Thr	Ala	Val	Tyr	Tyr	Cys	85	90	95	
Ala	Arg	Arg	Arg	Ile	Ala	Tyr	Gly	Tyr	Asp	Glu	Gly	His	Ala	Met	Asp	100	105	110	
Tyr	Trp	Gly	Gln	Gly	Thr	Leu	Val	Thr	Val	Ser	Ser					115	120		

(2) INFORMATION FOR SEQ ID NO: 41:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 124 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 41:

Gln	Val	Gln	Leu	Val	Gln	Ser	Gly	Ala	Glu	Val	Lys	Lys	Pro	Gly	Ala	1	5	10	15
Ser	Val	Lys	Val	Ser	Cys	Lys	Thr	Ser	Gly	Tyr	Thr	Phe	Thr	Glu	Tyr	20	25	30	
Thr	Ile	His	Trp	Val	Arg	Gln	Ala	Pro	Gly	Gln	Arg	Leu	Glu	Trp	Ile	35	40	45	
Gly	Gly	Ile	Asn	Pro	Asn	Asn	Gly	Ile	Pro	Asn	Tyr	Asn	Gln	Lys	Phe	50	55	60	
Lys	Gly	Arg	Val	Thr	Ile	Thr	Val	Asp	Thr	Ser	Ala	Ser	Thr	Ala	Tyr	65	70	75	80
Met	Glu	Leu	Ser	Ser	Leu	Arg	Ser	Glu	Asp	Thr	Ala	Val	Tyr	Tyr	Cys	85	90	95	

EP 0 953 639 A1

Ala Arg Arg Arg Ile Ala Tyr Gly Tyr Asp Glu Gly His Ala Met Asp
100 105 110

Tyr Trp Gly Gln Gly Thr Leu Val Thr Val Ser Ser
115 120

(2) INFORMATION FOR SEQ ID NO: 42:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 7731 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 42:

TTGAAGACGA	AAGGGCCTCG	TGATACGCCT	ATTTTATAG	GTAAATGTCA	TGATAATAAT	60
GGTTTCTTAG	ACGTCAGGTG	GCACTTTTCG	GGGAAATGTG	CGCGGAACCC	CTATTGTGTT	120
ATTTTCTAA	ATACATTCAA	ATATGTATCC	GCTCATGAGA	CAATAACCCT	GATAAATGCT	180
TCAATAATAT	TGAAAAAGGA	AGAGTATGAG	TATTCAACAT	TTCCGTGTGCG	CCCTTATTCC	240
CTTTTGTGCG	GCATTTTGCC	TTCTGTGTTT	TGCTCACCCA	GAAACGCTGG	TGAAAGTAAA	300
AGATGCTGAA	GATCAGTTGG	GTGCACGAGT	GGGTTACATC	GAACGCGATC	TCAACAGCGG	360
TAAGATCCTT	GAGAGTTTTC	GCCCCGAAGA	ACGTTTTCCT	ATGATGAGCA	CTTTTAAAGT	420
TCTGCTATGT	GGCGCGGTAT	TATCCCGTGT	TGACGCCGGG	CAAGAGCAAC	TCGGTCGCCG	480
CATACACTAT	TCTCAGAATG	ACTTGGTTGA	GTACTCACCA	GTCACAGAAA	AGCATCTTAC	540
GGATGGCATG	ACAGTAAGAG	AATTATGCAG	TGCTGCCATA	ACCATGAGTG	ATAACACTGC	600
GGCCAACTTA	CTTCTGACAA	CGATCGGAGG	ACCGAAGGAG	CTAACCCTTT	TTTTGCACAA	660
CATGGGGGAT	CATGTAAGTC	GCCTTGATCG	TTGGGAACCG	GAGCTGAATG	AAGCCATACC	720
AAACGACGAG	CGTGACACCA	CGATGCCTGC	AGCAATGGCA	ACAACGTTGC	GCAAACCTATT	780
AACTGGCGAA	CTACTTACTC	TAGCTTCCCG	GCAACAATTA	ATAGACTGGA	TGGAGGCGGA	840
TAAAGTTGCA	GGACCACTTC	TGCGCTCGGC	CCTTCCGGCT	GGCTGGTTTA	TTGCTGATAA	900
ATCTGGAGCC	GGTGAGCGTG	GGTCTCGCGG	TATCATTTGCA	GCACTGGGGC	CAGATGGTAA	960
GCCCTCCCGT	ATCGTAGTTA	TCTACACGAC	GGGGAGTCAG	GCAACTATGG	ATGAACGAAA	1020
TAGACAGATC	GCTGAGATAG	GTGCCTCACT	GATTAAGCAT	TGGTAACTGT	CAGACCAAGT	1080
TTACTCATAT	ATACTTTAGA	TTGATTTAAA	ACTTCATTTT	TAATTTAAAA	GGATCTAGGT	1140
GAAGATCCTT	TTTGATAATC	TCATGACCAA	AATCCCTTAA	CGTGAGTTT	CGTTCCACTG	1200
AGCGTCAGAC	CCCGTAGAAA	AGATCAAAGG	ATCTTCTTGA	GATCCTTTT	TTCTGCGCGT	1260
AATCTGCTGC	TTGCAAACAA	AAAAACCACC	GCTACCAGCG	GTGGTTTGTT	TGCCGGATCA	1320
AGAGCTACCA	ACTCTTTTTC	CGAAGGTAAC	TGGCTTCAGC	AGAGCGCAGA	TACCAAATAC	1380
TGTCCTTCTA	GTGTAGCCGT	AGTTAGGCCA	CCACTTCAAG	AACTCTGTAG	CACCGCCTAC	1440
ATACCTCGCT	CTGCTAATCC	TGTTACCAGT	GGCTGCTGCC	AGTGCGGATA	AGTCGTGTCT	1500

EP 0 953 639 A1

	TACCGGGTTG	GACTCAAGAC	GATAGTTACC	GGATAAGGCG	CAGCGGTCGG	GCTGAACGGG	1560
	GGGTTCGTGC	ACACAGCCCA	GCTTGGAGCG	AACGACCTAC	ACCGAACTGA	GATACCTACA	1620
5	GCGTGAGCTA	TGAGAAAGCG	CCACGCTTCC	CGAAGGGAGA	AAGGCGGACA	GGTATCCGGT	1680
	AAGCGGCAGG	GTCGGAACAG	GAGAGCGCAC	GAGGGAGCTT	CCAGGGGGAA	ACGCCTGGTA	1740
	TCTTTATAGT	CCTGTGCGGT	TTCCGCCACCT	CTGACTTGAG	CGTCGATTTT	TGTGATGCTC	1800
10	GTCAGGGGGG	CGGAGCCTAT	GGAAAAACGC	CAGCAACGCG	GCCTTTTTTAC	GGTTCCTGGC	1860
	CTTTTGCTGG	CCTTTTGCTC	ACATGTTCTT	TCCTGCGTTA	TCCCCTGATT	CTGTGGATAA	1920
	CCGTATTACC	GCCTTTGAGT	GAGCTGATAC	CGCTCGCCGC	AGCCGAACGA	CCGAGCGCAG	1980
	CGAGTCAGTG	AGCGAGGAAG	CGGAAGAGCG	CCTGATGCGG	TATTTTCTCC	TTACGCATCT	2040
15	GTGCGGTATT	TCACACCGCA	TATGGTGCAC	TCTCAGTACA	ATCTGCTCTG	ATGCCGCATA	2100
	GTTAAGCCAG	TATACACTCC	GCTATCGCTA	CGTGACTGGG	TCATGGCTGC	GCCCCGACAC	2160
	CCGCCAACAC	CCGCTGACGC	GCCCTGACGG	GCTTGTCTGC	TCCCGGCATC	CGCTTACAGA	2220
20	CAAGCTGTGA	CCGTCTCCGG	GAGCTGCATG	TGTCAGAGGT	TTTCACCGTC	ATCACCMAAA	2280
	CGCGCGAGGC	AGCATGCATC	TCAATTAGTC	AGCAACCATA	GTCCCGCCCC	TAACCTCCGCC	2340
	CATCCCGCCC	CTAATCCCGC	CCAGTTCCGC	CCATTCTCCG	CCCCATGGCT	GACTAATTTT	2400
	TTTTATTTAT	GCAGAGGCCG	AGGCCGCCCTC	GGCCTCTGAG	CTATTCCAGA	AGTAGTGAGG	2460
25	AGGCTTTTTT	GGAGGCCTAG	GCTTTTGCAA	AAAGCTAGCT	TACAGCTCAG	GGCTGCGATT	2520
	TCGCGCCAAA	CTTGACGGCA	ATCCTAGCGT	GAAGGCTGGT	AGGATTTTAT	CCCCGCTGCC	2580
	ATCATGGTTC	GACCATTGAA	CTGCATCGTC	GCCGTGTCCC	AAAATATGGG	GATTGGCAAG	2640
30	AACGGAGACC	TACCCTGGCC	TCCGCTCAGG	AACGAGTTCA	AGTACTTCCA	AAGAATGACC	2700
	ACAACCTCTT	CAGTGGAAGG	TAAACAGAAT	CTGGTGATTA	TGGGTAGGAA	AACCTGGTTC	2760
	TCCATTCCCTG	AGAAGAATCG	ACCTTTAAAG	GACAGAATTA	ATATAGTTCT	CAGTAGAGAA	2820
	CTCAAAGAAC	CACCACGAGG	AGCTCATTTT	CTTGCCAAAA	GTTTGGATGA	TGCCTTAAGA	2880
35	CTTATTGAAC	AACCGGAATT	GGCAAGTAAA	GTAGACATGG	TTTGATAGT	CGGAGGCAGT	2940
	TCTGTTTACC	AGGAAGCCAT	GAATCAACCA	GGCCACCTCA	GACTCTTTGT	GACAAGGATC	3000
	ATGCAGGAAT	TTGAAAGTGA	CACGTTTTTC	CCAGAAATTG	ATTTGGGGAA	ATATAAACTT	3060
40	CTCCCAGAAT	ACCCAGGCGT	CCTCTCTGAG	GTCCAGGAGG	AAAAAGGCAT	CAAGTATAAG	3120
	TTTGAAGTCT	ACGAGAAGAA	AGACTAACAG	GAAGATGCTT	TCAAGTTCTC	TGCTCCCCCTC	3180
	CTAAAGCTAT	GCATTTTTAT	AAGACCATGG	GACTTTTGCT	GGCTTTAGAT	CTTTGTGAAG	3240
45	GAACCTTACT	TCTGTGGTGT	GACATAATTG	GACAACTAC	CTACAGAGAT	TTAAAGCTCT	3300
	AAGGTAAATA	TAAATTTTTT	AAGTGATATA	TGTGTTAAAC	TACTGATTCT	AATTGTTTTGT	3360
	GTATTTTAGA	TTCCAACCTA	TGGAACGAT	GAATGGGAGC	AGTGGTGGAA	TGCCTTTAAT	3420
	GAGGAAAACC	TGTTTTGCTC	AGAAGAAATG	CCATCTAGTG	ATGATGAGGC	TACTGCTGAC	3480
50	TCTCAACATT	CTACTCCTCC	AAAAAAGAAG	AGAAAGGTAG	AAGACCCCAA	GGACTTTCCT	3540
	TCAGAATTGC	TAAGTTTTTT	GAGTCATGCT	GTGTTTAGTA	ATAGAACTCT	TGCTTGCTTT	3600

55

EP 0 953 639 A1

	GCTATTTACA	CCACAAAGGA	AAAAGCTGCA	CTGCTATACA	AGAAAAATTAT	GGAAAAATAT	3660
	TCTGTAACCT	TTATAAGTAG	GCATAACAGT	TATAATCATA	ACATACTGTT	TTTTCTTACT	3720
5	CCACACAGGC	ATAGAGTGTC	TGCTATTAAT	AACTATGCTC	AAAAATTGTG	TACCTTTAGC	3780
	TTTTTAATTT	GTAAAGGGGT	TAATAAGGAA	TATTTGATGT	ATAGTGCCTT	GACTAGAGAT	3840
	CATAATCAGC	CATACCACAT	TTGTAGAGGT	TTTACTTGCT	TTAAAAAACC	TCCCACACCT	3900
10	CCCCCTGAAC	CTGAAACATA	AAATGAATGC	AATTGTTGTT	GTAACTTGT	TTATTGCAGC	3960
	TTATAATGGT	TACAAATAAA	GCAATAGCAT	CACAAATTTT	ACAAATAAAG	CATTTTTTTT	4020
	ACTGCATTCT	AGTTGTGGTT	TGTCCAAACT	CATCAATGTA	TCTTATCATG	TCTGGATCTA	4080
	ATAAAAGATA	TTTATTTTCA	TTAGATATGT	GTGTTGGTTT	TTTGTGTGCA	GTGCCTCTAT	4140
15	CTGGAGGCCA	GGTAGGGCTG	GCCTTGGGGG	AGGGGGAGGC	CAGAATGACT	CCAAGAGCTA	4200
	CAGGAAGGCA	GGTCAGAGAC	CCCACTGGAC	AAACAGTGGC	TGGACTCTGC	ACCATAACAC	4260
	ACAATCAACA	GGGGAGTGAG	CTGGAAATTT	GCTAGCGAAT	TCCAGCACAC	TGGCGGCCGT	4320
20	TACTAGTTAT	TAATAGTAAT	CAATTACGGG	GTCATTAGTT	CATAGCCCAT	ATATGGAGTT	4380
	CCGCGTTACA	TAACTTACGG	TAAATGGCCC	GCCTGGCTGA	CCGCCCAACG	ACCCCCGCCC	4440
	ATTGACGTCA	ATAATGACGT	ATGTTCCCAT	AGTAACGCCA	ATAGGGACTT	TCCATTGACG	4500
	TCAATGGGTG	GAGTATTTAC	GGTAAACTGC	CCACTTGGCA	GTACATCAAG	TGTATCATAT	4560
25	GCCAAGTACG	CCCCCTATTG	ACGTCAATGA	CGGTAAATGG	CCCGCCTGGC	ATTATGCCCA	4620
	GTACATGACC	TTATGGGACT	TTCTTACTTG	GCAGTACATC	TACGTATTAG	TCATCGCTAT	4680
	TACCATGGTG	ATGCGGTTTT	GGCAGTACAT	CAATGGGCGT	GGATAGCGGT	TTGACTCACG	4740
30	GGGATTTCCA	AGTCTCCACC	CCATTGACGT	CAATGGGAGT	TTGTTTTGGC	ACCAAAATCA	4800
	ACGGGACTTT	CCAAAATGTC	GTAACAATC	CGCCCCATTG	ACGCAAATGG	GCGGTAGGCG	4860
	TGTACGGTGG	GAGGTCTATA	TAAGCAGAGC	TCGTTTAGTG	AACCGTCAGA	TCGCCTGGAG	4920
35	ACGCCATCCA	CGCTGTTTTG	ACCTCCATAG	AAGACACCGG	GACCGATCCA	GCCTCCGCGG	4980
	CCGGGAACGG	TGCATTGGAA	CGCGGATTCC	CCGTGCCAAG	AGTGACGTAA	GTACCGCCTA	5040
	TAGAGTCTAT	AGGCCACCCC	CCTTGGCTTC	TTATGCATGC	TATACTGTTT	TTGGCTTGGG	5100
	GTCTATACAC	CCCCGCTTCC	TCATGTTATA	GGTGATGGTA	TAGCTTAGCC	TATAGGTGTG	5160
40	GGTTATTGAC	CATTATTGAC	CACTCCCCCTA	TTGGTGACGA	TACTTTCCAT	TACTAATCCA	5220
	TAACATGGCT	CTTTGCCACA	ACTCTCTTTA	TTGGCTATAT	GCCAATACAC	TGTCCTTCAG	5280
	AGACTGACAC	GGACTCTGTA	TTTTTACAGG	ATGGGGTCTC	ATTTATTATT	TACAAATTCA	5340
45	CATATACAAC	ACCACCGTCC	CCAGTGCCCG	CAGTTTTTTAT	TAAACATAAC	GTGGGATCTC	5400
	CACGCGAATC	TCGGGTACGT	GTTCCGGACA	TGGGCTCTTC	TCCGGTAGCG	GCGGAGCTTC	5460
	TACATCCGAG	CCCTGCTCCC	ATGCCTCCAG	CGACTCATGG	TCGCTCGGCA	GCTCCTTGCT	5520
	CCTAACAGTG	GAGGCCAGAC	TTAGGCACAG	CACGATGCCC	ACCACCACCA	GTGTGCCGCA	5580
50	CAAGGCCGTG	GCGGTAGGGT	ATGTGTCTGA	AAATGAGCTC	GGGGAGCGGG	CTTGCAACGC	5640
	TGACGCATTT	GGAAGACTTA	AGGCAGCGGC	AGAAGAAGAT	GCAGGCAGCT	GAGTTGTTGT	5700
55							

EP 0 953 639 A1

	GTTCTGATAA	GAGTCAGAGG	TAACCTCCCGT	TGCGGTGCTG	TTAACGGTGG	AGGGCAGTGT	5760
	AGTCTGAGCA	GTACTCGTTG	CTGCCGCGCG	CGCCACCAGA	CATAATAGCT	GACAGACTAA	5820
5	CAGACTGTTC	CTTTCCATGG	GTCTTTTCTG	CAGTCACCGT	CCTTGACACG	CGTCTCGGGA	5880
	AGCTTGCCGC	CACCATGGAC	TGGACCTGGC	GCGTGTTTTG	CCTGCTCGCC	GTGGCTCCTG	5940
	GGGCCCACAG	CCAGGTGCAA	CTGGTGCACT	CCGGCGCCGA	AGTGAAGAAA	CCCGGTGCTT	6000
10	CCGTGAAAGT	CAGCTGTAAA	ACTAGTAGAT	ACACCTTCAC	TGAATACACC	ATACACTGGG	6060
	TTAGACAGGC	CCCTGGCCAA	AGGCTGGAGT	GGATAGGAGG	TATTAATCCT	AACAATGGTA	6120
	TTCTTAACCTA	CAACCAGAAG	TTCAAGGGCC	GGGCCACCTT	GACCGTAGGC	AAGTCTGCCA	6180
	GCACCGCCTA	CATGGAAGT	TCCAGCCTGC	GCTCCGAGGA	CACTGCAGTC	TACTACTGCG	6240
15	CCAGAAGAAG	AATCGCCTAT	GGTTACGACG	AGGGCCATGC	TATGGACTAC	TGGGGTCAAG	6300
	GAACCTTGT	CACCGTCTCC	TCAGGTGAGT	GGATCCTCTG	CGCCTGGGCC	CAGCTCTGTC	6360
	CCACACCGCG	GTCACATGGC	ACCACCTCTC	TTGCAGCCTC	CACCAAGGGC	CCATCGGTCT	6420
20	TCCCCCTGGC	ACCTTCCTCC	AAGAGCACCT	CTGGGGGCAC	AGCGGCCCTG	GGCTGCCTGG	6480
	TCAAGGACTA	CTTCCCCGAA	CCGGTGACGG	TGTCGTGGAA	CTCAGGCGCC	CTGACCAGCG	6540
	GCGTGACAC	CTTCCCGGCT	GTCCTACAGT	CCTCAGGACT	CTACTCCCTC	AGCAGCGTGG	6600
25	TGACCGTGCC	CTCCAGCAGC	TTGGGCACCC	AGACCTACAT	CTGCAACGTG	AATCACAAGC	6660
	CCAGCAACAC	CAAGGTGGAC	AAGAAAGTTG	AGCCCAAATC	TTGTGACAAA	ACTCACACAT	6720
	GCCCACCGTG	CCCAGCACCT	GAACCTCTGG	GGGGACCGTC	AGTCTTCCTC	TTCCCCCCTA	6780
	AACCCAAGGA	CACCCTCATG	ATCTCCCGGA	CCCCTGAGGT	CACATGCGTG	GTGGTGGACG	6840
30	TGAGCCACGA	AGACCCTGAG	GTCAAGTTCA	ACTGGTACGT	GGACGGCGTG	GAGGTGCATA	6900
	ATGCCAAGAC	AAAGCCGCGG	GAGGAGCAGT	ACAACAGCAC	GTACCGGGTG	GTCAGCGTCC	6960
	TCACCGTCCT	GCACCAGGAC	TGGCTGAATG	GCAAGGAGTA	CAAGTGCAAG	GTCTCCAACA	7020
35	AAGCCCTCCC	AGCCCCCATC	GAGAAAACCA	TCTCCAAAGC	CAAAGGGCAG	CCCCGAGAAC	7080
	CACAGGTGTA	CACCCTGCCC	CCATCCCGGG	AGGAGATGAC	CAAGAACCAG	GTCAGCCTGA	7140
	CCTGCCTGGT	CAAAGGCTTC	TATCCCAGCG	ACATCGCCGT	GGAGTGGGAG	AGCAATGGGC	7200
	AGCCGAGAGAA	CAACTACAAG	ACCACGCCTC	CCGTGCTGGA	CTCCGACGGC	TCCTTCTTCC	7260
40	TCTACAGCAA	GCTCACCGTG	GACAAGAGCA	GGTGGCAGCA	GGGGAACGTC	TTCTCATGCT	7320
	CCGTGATGCA	TGAGGCTCTG	CACAACCACT	ACACGCAGAA	GAGCCTCTCC	CTGTCTCCGG	7380
	GTAAATGAGT	GCGACGGCCG	GCAAGCCCCG	CTCCCCGGGC	TCTCGCGGTC	GCACGAGGAT	7440
45	GCTTGGCACG	TACCCCTGT	ACATACTTCC	CGGGCGCCCA	GCATGGAAAT	AAAGCACCGG	7500
	ATCTAATAAA	AGATATTTAT	TTTCATTAGA	TATGTGTGTT	GGTTTTTTGT	GTGCAGTGCC	7560
	TCTATCTGGA	GGCCAGGTAG	GGCTGGCCTT	GGGGGAGGGG	GAGGCCAGAA	TGACTCCAAG	7620
	AGCTACAGGA	AGGCAGGTCA	GAGACCCAC	TGGACAAACA	GTGGCTGGAC	TCTGCACCAT	7680
50	AACACACAAT	CAACAGGGGA	GTGAGCTGGA	AATTTGCTAG	CGAATTAATT	C	7731

(2) INFORMATION FOR SEQ ID NO: 43:

55

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 472 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 43:

Met	Asp	Trp	Thr	Trp	Arg	Val	Phe	Cys	Leu	Leu	Ala	Val	Ala	Pro	Gly	1	5	10	15
Ala	His	Ser	Gln	Val	Gln	Leu	Val	Gln	Ser	Gly	Ala	Glu	Val	Lys	Lys	20	25	30	
Pro	Gly	Ala	Ser	Val	Lys	Val	Ser	Cys	Lys	Thr	Ser	Arg	Tyr	Thr	Phe	35	40	45	
Thr	Glu	Tyr	Thr	Ile	His	Trp	Val	Arg	Gln	Ala	Pro	Gly	Gln	Arg	Leu	50	55	60	
Glu	Trp	Ile	Gly	Gly	Ile	Asn	Pro	Asn	Asn	Gly	Ile	Pro	Asn	Tyr	Asn	65	70	75	80
Gln	Lys	Phe	Lys	Gly	Arg	Ala	Thr	Leu	Thr	Val	Gly	Lys	Ser	Ala	Ser	85	90	95	
Thr	Ala	Tyr	Met	Glu	Leu	Ser	Ser	Leu	Arg	Ser	Glu	Asp	Thr	Ala	Val	100	105	110	
Tyr	Tyr	Cys	Ala	Arg	Arg	Arg	Ile	Ala	Tyr	Gly	Tyr	Asp	Glu	Gly	His	115	120	125	
Ala	Met	Asp	Tyr	Trp	Gly	Gln	Gly	Thr	Leu	Val	Thr	Val	Ser	Ser	Ser	130	135	140	
Thr	Lys	Gly	Pro	Ser	Val	Phe	Pro	Leu	Ala	Pro	Ser	Ser	Lys	Ser	Thr	145	150	155	160
Ser	Gly	Gly	Thr	Ala	Ala	Leu	Gly	Cys	Leu	Val	Lys	Asp	Tyr	Phe	Pro	165	170	175	
Glu	Pro	Val	Thr	Val	Ser	Trp	Asn	Ser	Gly	Ala	Leu	Thr	Ser	Gly	Val	180	185	190	
His	Thr	Phe	Pro	Ala	Val	Leu	Gln	Ser	Ser	Gly	Leu	Tyr	Ser	Leu	Ser	195	200	205	
Ser	Val	Val	Thr	Val	Pro	Ser	Ser	Ser	Leu	Gly	Thr	Gln	Thr	Tyr	Ile	210	215	220	
Cys	Asn	Val	Asn	His	Lys	Pro	Ser	Asn	Thr	Lys	Val	Asp	Lys	Lys	Val	225	230	235	240
Glu	Pro	Lys	Ser	Cys	Asp	Lys	Thr	His	Thr	Cys	Pro	Pro	Cys	Pro	Ala	245	250	255	
Pro	Glu	Leu	Leu	Gly	Gly	Pro	Ser	Val	Phe	Leu	Phe	Pro	Pro	Lys	Pro	260	265	270	
Lys	Asp	Thr	Leu	Met	Ile	Ser	Arg	Thr	Pro	Glu	Val	Thr	Cys	Val	Val	275	280	285	
Val	Asp	Val	Ser	His	Glu	Asp	Pro	Glu	Val	Lys	Phe	Asn	Trp	Tyr	Val	290	295	300	

Asp Gly Val Glu Val His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln
 305 310 315 320
 Tyr Asn Ser Thr Tyr Arg Val Val Ser Val Leu Thr Val Leu His Gln
 325 330 335
 Asp Trp Leu Asn Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Ala
 340 345 350
 Leu Pro Ala Pro Ile Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro
 355 360 365
 Arg Glu Pro Gln Val Tyr Thr Leu Pro Pro Ser Arg Glu Glu Met Thr
 370 375 380
 Lys Asn Gln Val Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser
 385 390 395 400
 Asp Ile Ala Val Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr
 405 410 415
 Lys Thr Thr Pro Pro Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr
 420 425 430
 Ser Lys Leu Thr Val Asp Lys Ser Arg Trp Gln Gln Gly Asn Val Phe
 435 440 445
 Ser Cys Ser Val Met His Glu Ala Leu His Asn His Tyr Thr Gln Lys
 450 455 460
 Ser Leu Ser Leu Ser Pro Gly Lys
 465 470

(2) INFORMATION FOR SEQ ID NO: 44:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 25 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 44:

ACCGTCTCCT CAGGTGAGTG GATCC

25

(2) INFORMATION FOR SEQ ID NO: 45:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 14 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 45:

CCTCTCTTGC AGCC

14

(2) INFORMATION FOR SEQ ID NO: 46:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 14 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 46:

CCTCTCTTGC AGCC

(2) INFORMATION FOR SEQ ID NO: 47:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 4 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 47:

Thr Val Ser Ser
1

(2) INFORMATION FOR SEQ ID NO: 48:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 4 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 48:

Ser Thr Lys Gly
1

(2) INFORMATION FOR SEQ ID NO: 49:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 27 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 49:

ACCGTCTCCT CAGCCTCCAC CAAGGGC

(2) INFORMATION FOR SEQ ID NO: 50:

5 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 8 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

10 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 50:
Thr Val Ser Ser Ser Thr Lys Gly
1 5

15 (2) INFORMATION FOR SEQ ID NO: 51:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 27 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

20 (ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 51:
25 ACCGTCTCCT CAGCCTCCAC CAAGGGC 27

(2) INFORMATION FOR SEQ ID NO: 52:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 9 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

30 (ii) MOLECULE TYPE: peptide

35 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 52:
Thr Val Ser Ser Ala Ser Thr Lys Gly
1 5

40 (2) INFORMATION FOR SEQ ID NO: 53:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

45 (ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 53:
50 GAAATAAAAC GTGAGTGGAT CC 22

(2) INFORMATION FOR SEQ ID NO: 54:

55

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 27 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 54:

CTTCTTTTCCT CAGGAACTGT GGCTGCA

27

(2) INFORMATION FOR SEQ ID NO: 55:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 4 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 55:

Thr Val Ala Ala
1

(2) INFORMATION FOR SEQ ID NO: 56:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 56:

GAAATAAAAC GAACTGTGGC TGCA

24

(2) INFORMATION FOR SEQ ID NO: 57:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 7 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 57:

Glu Ile Lys Thr Val Ala Ala
1 5

(2) INFORMATION FOR SEQ ID NO: 58:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 24 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 58:

GAAATAAAAC GAACTGTGGC TGCA

24

(2) INFORMATION FOR SEQ ID NO: 59:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 8 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 59:

Glu Ile Lys Arg Thr Val Ala Ala
 1 5

(2) INFORMATION FOR SEQ ID NO: 60:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 20 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 60:

Met Asp Ser Gln Ala Gln Val Leu Met Leu Leu Leu Leu Trp Val Ser
 1 5 10 15

Gly Thr Cys Gly
 20

(2) INFORMATION FOR SEQ ID NO: 61:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 19 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 61:

Met Gly Trp Ser Trp Val Phe Leu Phe Leu Leu Ser Gly Thr Ala Gly

1 5 10 15

Val Leu Ser

5

(2) INFORMATION FOR SEQ ID NO: 62:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 9 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

10

(ii) MOLECULE TYPE: DNA (genomic)

15

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 62:

GCCGCCACC

9

(2) INFORMATION FOR SEQ ID NO: 63:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 37 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

20

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "PRIMER"

25

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 63:

CAGAAAGCTT GCCGCCACCA TGGATTCACA GGCCCAG

30

37

(2) INFORMATION FOR SEQ ID NO: 64:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 6 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

35

(ii) MOLECULE TYPE: peptide

40

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 64:

Met Asp Ser Gln Ala Gln
1 5

45

(2) INFORMATION FOR SEQ ID NO: 65:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 35 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

50

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "PRIMER"

55

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 65:

CCGAGGATCC ACTCACGTTT CAGCTCCAGC TTGGT

35

(2) INFORMATION FOR SEQ ID NO: 66:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 37 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid
 - (A) DESCRIPTION: /desc = "PRIMER"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 66:

CAGAAAGCTT GCCGCCACCA TGGGATGGAG CTGGGTC

37

(2) INFORMATION FOR SEQ ID NO: 67:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 6 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 67:

Met Gly Trp Ser Trp Val
1 5

(2) INFORMATION FOR SEQ ID NO: 68:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 35 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid
 - (A) DESCRIPTION: /desc = "PRIMER"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 68:

CCGAGGATCC ACTCACCTGA GGAGACGGTG ACTGA

35

(2) INFORMATION FOR SEQ ID NO: 69:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 36 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid
 - (A) DESCRIPTION: /desc = "PRIMER"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 69:

GTTCATCACAA TGTCTCCGGA GGAACCTGGA ACCCAG

36

(2) INFORMATION FOR SEQ ID NO: 70:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 29 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "PRIMER"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 70:

CTCCGGAGAC ATTGTGATGA CCCAATCTC

29

(2) INFORMATION FOR SEQ ID NO: 71:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 29 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "PRIMER"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 71:

CTCCGGAGAC ATTGTGATGA CCCAATCTC

29

(2) INFORMATION FOR SEQ ID NO: 72:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 72 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "PRIMER"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 72:

CAGTCAGAGC CTTTATATT CTAGAAATCA AAAGAACTAC TTGGCCTGGT ATCAGCAGAA

60

ACCAGGACAG CC

72

(2) INFORMATION FOR SEQ ID NO: 73:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 44 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "PRIMER"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 73:

ACCCCAGATT CCCTAGTGCT AGCCCAAAG ATGAGGAGTT TGGG

(2) INFORMATION FOR SEQ ID NO: 74:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 67 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "PRIMER"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 74:

TAGCACTAGG GAATCTGGGG TACCTGATAG GTTCAGTGGC AGTGGGTTTG GGACAGACTT
CACCCCTC

(2) INFORMATION FOR SEQ ID NO: 75:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 53 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "PRIMER"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 75:

GTCCCTTGTC CGAACGTGAG CGGATAGCTA AAATATTGCT GACAGTAATA AAC

(2) INFORMATION FOR SEQ ID NO: 76:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 33 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "PRIMER"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 76:

GCTCACGTTT GGACAAGGGA CCAAGGTGGA AAT

(2) INFORMATION FOR SEQ ID NO: 77:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 72 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "PRIMER"

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 77:

CAGTCAGAGC CTTTATATATT CTAGAAATCA AAAGAACTAC TTGGCCTGGT TCCAGCAGAA 60
ACCAGGACAG CC 72

- (2) INFORMATION FOR SEQ ID NO: 78:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 57 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "PRIMER"

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 78:

GTCCCTTGTC CGAACGTGAG CGGATAGCTA AAATATTGCT GACAGTCATA AACTGCC 57

- (2) INFORMATION FOR SEQ ID NO: 79:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 34 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "PRIMER"

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 79:

CCCAAACCTCC TCATCTATTG GGCTAGCACT AGGG 34

- (2) INFORMATION FOR SEQ ID NO: 80:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 34 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "PRIMER"

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 80:

CCCTAGTGCT AGCCCAATAG ATGAGGAGTT TGGG

34

(2) INFORMATION FOR SEQ ID NO: 81:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "PRIMER"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 81:

TACGCAAACC GCCTCTC

17

(2) INFORMATION FOR SEQ ID NO: 82:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "PRIMER"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 82:

GAGTGCACCA TATGCGGT

18

(2) INFORMATION FOR SEQ ID NO: 83:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 16 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "PRIMER"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 83:

AACAGCTATG ACCATG

16

(2) INFORMATION FOR SEQ ID NO: 84:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "PRIMER"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 84:

GTTTTCCCAG TCACGAC

17

5 (2) INFORMATION FOR SEQ ID NO: 85:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 47 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

10

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "PRIMER"

15 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 85:

GTGTATTTCAG TGAAGGTGTA TCTACTAGTT TTACAGCTGA CTTTCAC

47

(2) INFORMATION FOR SEQ ID NO: 86:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 53 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

20

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "PRIMER"

25

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 86:

TAGTAGATAC ACCTTCACTG AATACACCAT ACACTGGGTT AGACAGGCCCTG

53

30 (2) INFORMATION FOR SEQ ID NO: 87:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 71 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

35

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "PRIMER"

40 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 87:

CCCTTGAAC TCTGGTTGTA GTTAGGAATA CCATTGTTAG GATTAATACC TCCTATCCAC

60

TCCAGCCTTT G

71

45 (2) INFORMATION FOR SEQ ID NO: 88:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 71 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

50

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "PRIMER"

55

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 88:

5 TAACTACAAC CAGAAGTTCA AGGGCCGGGC CACCTTGACC GTAGGCAAGT CTGCCAGCAC 60
CGCCTACATG G 71

(2) INFORMATION FOR SEQ ID NO: 89:

(i) SEQUENCE CHARACTERISTICS:
10 (A) LENGTH: 63 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
15 (A) DESCRIPTION: /desc = "PRIMER"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 89:

20 GCATGGCCCT CGTCGTAACC ATAGGCGATT CTTCTTCTGG CGCAGTAGTA GACTGCAGTG 60
TCC 63

(2) INFORMATION FOR SEQ ID NO: 90:

(i) SEQUENCE CHARACTERISTICS:
25 (A) LENGTH: 48 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
30 (A) DESCRIPTION: /desc = "PRIMER"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 90:

35 CTATGGTTAC GACGAGGGCC ATGCTATGGA CTACTGGGGT CAAGGAAC 48

(2) INFORMATION FOR SEQ ID NO: 91:

(i) SEQUENCE CHARACTERISTICS:
40 (A) LENGTH: 71 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
45 (A) DESCRIPTION: /desc = "PRIMER"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 91:

50 TAACTACAAC CAGAAGTTCA AGGGCCGGGT CACCATCACC GTAGACACCT CTGCCAGCAC 60
CGCCTACATG G 71

(2) INFORMATION FOR SEQ ID NO: 92:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 27 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "PRIMER"

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 92:

GGACACTGCA GTCTACTTCT GCGCCAG

- (2) INFORMATION FOR SEQ ID NO: 93:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "PRIMER"

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 93:

TACGCAAACC GCCTCTC

- (2) INFORMATION FOR SEQ ID NO: 94:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "PRIMER"

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 94:

GAGTGCACCA TATGCGGT

- (2) INFORMATION FOR SEQ ID NO: 95:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 76 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "PRIMER"

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 95:

CCTTTGGCCA GGGGCCTGTC TAACCCAGTG TATGGTGTAT TCAGTGAAGG TGCTATCCAC

TAGTTTCCAC TAGTTT

76

(2) INFORMATION FOR SEQ ID NO: 96:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 28 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "PRIMER"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 96:

GTCACCGTCC TTGACACGCG TCTCGGGA

28

(2) INFORMATION FOR SEQ ID NO: 97:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "PRIMER"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 97:

TTGGAGGAGG GTGCCAG

17

(2) INFORMATION FOR SEQ ID NO: 98:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "PRIMER"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 98:

GAGACATTGT GACCCAATCT CC

22

(2) INFORMATION FOR SEQ ID NO: 99:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "PRIMER"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 99:

GACAGTCATA AACTGCCACA TCTTC

25

(2) INFORMATION FOR SEQ ID NO: 100:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 23 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "PRIMER"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 100:

TTGACACGCG TCTCGGGAAG CTT

23

(2) INFORMATION FOR SEQ ID NO: 101:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "PRIMER"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 101:

GGCGCAGAGG ATCCACTCAC CT

22

Claims

1. An antibody protein having the complementary determining regions of the monoclonal antibody F19 (ATCC Accession No. HB 8269), said antibody protein specifically binding to fibroblast activation protein, characterized in that it has framework modifications resulting in the improved producibility in host cells as compared to a chimeric antibody having the variable regions of F19 and foreign constant regions.

2. An antibody protein characterised in that it has a variable light chain region and a variable heavy chain region according to claim 1, each joined to a human constant region.

3. The antibody protein of claim 2, wherein said human constant region of the light chain is a human kappa constant region.

4. The antibody protein of claim 2, wherein said human constant region of the heavy chain is a human gamma-1 constant region.

5. An antibody protein according to any one of claims 1 to 4, characterised in that its expression levels in crude media samples as determined by ELISA and/or purified antibody yields exceed the expression levels and/or purification yields of the chimeric antibodies without framework modifications by at least a factor of 10.

6. An antibody protein according to any one of claims 1 to 4, characterised in that its expression levels in crude media samples as determined by ELISA and/or purified antibody yields exceed the expression levels and/or purification yields of the chimeric antibodies without framework modifications by at least a factor of 20.

7. An antibody protein according to any one of claims 1 to 4, characterised in that its expression levels in crude media samples as determined by ELISA and/or purified antibody yields exceed the expression levels and/or purification yields of the chimeric antibodies without framework modifications by at least a factor of 100.

8. An antibody protein according to any one of claims 1 to 7, characterised in that it displays improved producibility in eucaryotic cells.

9. The antibody protein according to claim 8 wherein said eucaryotic cell is a chinese hamster ovary cell (CHO cell).

10. An antibody protein according to any one of claims 1 to 9, wherein the amino acid in Kabat position 87 of the light chain region is not asparagine.

11. The antibody protein of claim 10, wherein the amino acid in Kabat position 87 of the light chain region is selected from aromatic or aliphatic amino acids.

12. The antibody protein of claim 11, wherein said aromatic amino acid in Kabat position 87 of the light chain region is a tyrosine or phenylalanine.

13. The antibody protein according to any one of claims 1 to 12, wherein the amino acid in Kabat position 36 of the light chain region is selected from aromatic amino acids.

14. An antibody protein according to any one of claims 1 to 13 that contains the variable region of the light chain as set forth in SEQ ID NO: 2.

15. An antibody protein of claim 14 characterised in that the variable region of the light chain is encoded by a nucleotide sequence as set forth in SEQ ID NO: 1.

16. An antibody protein according to any one of claims 1 to 13 that contains the variable region of the light chain as set forth in SEQ ID NO: 6.

17. An antibody protein of claim 16 characterised in that the variable region of the light chain is encoded by a nucleotide sequence as set forth in SEQ ID NO: 5.

18. An antibody protein according to any one of claims 1 to 17 containing a variable region of the heavy chain as set forth in any one of SEQ ID NOs: 8, 10, 12, 14.

19. An antibody protein according to claim 18 characterised in that the variable region of the heavy chain is encoded by a nucleotide sequence as set forth in SEQ ID NOs: 7, 9, 11, 13.

20. An antibody protein according to any one of claims 1 to 14 containing the variable region of the light chain as set forth in SEQ ID NO: 2 and the variable region of the heavy chain as set forth in SEQ ID NOs: 12.

21. The antibody protein of claim 20 characterised in that the variable region of the the light chain is encoded by a nucleotide sequence as set forth in SEQ ID NO: 1 and the variable region of the heavy chain is encoded by a nucleotide sequence as set forth in SEQ ID NO: 11.

23. An antibody protein according to any one of claims 1 to 13 containing the variable region of the light chain as set forth in SEQ ID NO: 2 and the variable region of the heavy chain as set forth in SEQ ID NOs: 8.

24. The antibody protein of claim 23 characterised in that the variable region of the the light chain is encoded by a nucleotide sequence as set forth in SEQ ID NO: 1 and the variable region of the heavy chain is encoded by a nucleotide sequence as set forth in SEQ ID NO: 7.

25. A nucleotide sequence encoding an antibody protein according to any one of claims 1 to 24.

26. A recombinant DNA vector that contains a nucleotide sequence of claim 25.

27. The recombinant DNA vector of claim 26, said vector being an expression vector.

28. A host cell carrying a vector according to claims 26 or 27.

29. The host cell of claim 28, wherein said host cell is a eucaryotic cell.

30. The host cell of claim 29, wherein said eucaryotic host cell is a mammalian cell.

31. The host cell of claim 30, wherein said host cell is a CHO or a COS cell.

32. A method of producing antibody proteins according to any one of claims 1 to 24, said method comprising the steps of:

- (a) cultivating a host cell according to any one of claims 23 to 26 under conditions where said antibody protein is expressed by said host cell, and
- (b) isolating said antibody protein.

33. The method of claim 32, wherein said host cell is a mammalian cell, preferably a CHO or COS cell.

34. The method of claim 32 or 33, wherein said host cell is cotransfected with two plasmids carrying the expression units for light and heavy chains respectively.

35. An antibody protein according to any one of claims 1 to 24, wherein said antibody protein is conjugated to a therapeutic agent.

36. The antibody protein of claim 35, wherein said therapeutic agent is a therapeutic agent selected from the group consisting of radioisotopes, toxins, toxoids, inflammatory agents and chemotherapeutic agents.

37. The antibody protein of claim 36, wherein said radioisotopes are β -emitting radioisotopes.

38. The antibody protein of claim 37, wherein said radioisotopes are selected from the group consisting of ^{186}Re , ^{188}Re , ^{131}I and ^{90}Y .

39. An antibody protein according to any one of claims 1 to 24, characterised in that it is labeled.

40. The antibody protein of claim 39, wherein said label is a detectable marker.

41. The antibody protein of claim 40, wherein the detectable marker is a detectable marker selected from the group consisting of enzymes, dyes, radioisotopes, and biotin.

42. An antibody protein according to any one of claims 1 to 24 conjugated to an imageable agent.

43. The antibody protein of claim 42, wherein the imageable agent is a radioisotope.

44. The antibody protein of claim 43, wherein said radioisotopes are gamma-emitting radioisotopes??.

45. The antibody protein of claim 44, wherein said radioisotopes is ^{125}I .

46. A pharmaceutical composition containing an antibody protein according to any one of claims 1 to 24 and a pharmaceutically acceptable carrier useful for treating tumors, wherein said tumors are associated with activated stromal fibroblasts.

47. A pharmaceutical composition containing an antibody protein according to any one of claims 35 to 38 and a pharmaceutically acceptable carrier useful for treating tumors, wherein said tumors are associated with activated stromal fibroblasts.

48. A pharmaceutical composition containing an antibody protein according to any one of claims 42 to 45 and a pharmaceutically acceptable carrier useful for imaging the presence of activated stromal fibroblasts in a healing wound, inflamed skin or a tumor, in a human patient.

49. The pharmaceutical composition of claims 46 to 48, wherein said tumors are tumors selected from the cancer group consisting of colorectal cancers, non-small cell lung cancers, breast cancers, head and neck cancer, ovarian cancers, lung cancers, bladder cancers, pancreatic cancers and metastatic cancers of the brain.

50. Use of an antibody protein according to anyone of claims 1 to 24 for the treatment of cancer.

51. Use of an antibody protein according to anyone of claims 35 to 38 for the treatment of cancer.

52. Use of an antibody protein according to anyone of claims 42 to 45 for imaging activated activated stromal fibroblasts.

53. Use of an antibody protein according to anyone of claims 39 to 41 for detecting the presence of activated stromal fibroblasts in a sample.

54. A method of treating tumors, wherein the tumor is associated with activated stromal fibroblasts capable of specifically forming a complex with antibody proteins according to any one of claims 1 to 24 or 35 to 38, which comprises contacting the tumor with an amount of said antibody proteins effective to treat the tumor.

55. The method of claim 54, wherein the tumor is a tumor having cancer cells selected from the cancer group consisting of colorectal cancers, non-small cell lung cancers, breast cancers, head and neck cancer, ovarian cancers, lung cancers, bladder cancers, pancreatic cancers and metastatic cancers of the brain.

56. The method of claim 54, wherein the contacting is effected in vitro.

57. The method of claim 54, wherein the contacting is effected in vivo.

58. A method of detecting the presence of activated stromal fibroblasts in wound healing, inflammation or a tumor, characterised in that

(a) a sample, possibly containing activated stromal fibroblasts, is contacted with an antibody protein according to any one of claims 1 to 24 or 39 to 41 under conditions suitable for the formation of a complex between said antibody and antigen,

(b) detecting the presence of said complex, thereby detecting the presence of activated stromal fibroblasts in wound healing, inflammation or a tumor.

59. The method of claim 58, wherein the tumor is a tumor having cancer cells selected from the cancer group consisting of colorectal cancers, non-small cell lung cancers, breast cancers, head and neck cancer, ovarian cancers, lung cancers, bladder cancers, pancreatic cancers and metastatic cancers of the brain.

60. The method of claim 58 or 59, wherein the antibody protein is a protein according to any one of claims 39 to 41.

61. A method of imaging the presence of activated stromal fibroblasts in a healing wound, inflamed skin or a tumor, in a human patient, characterised in that

(a) an antibody protein according to any one of claims 1 to 24 conjugated to an imageable agent is administered to a human patient under conditions suitable for the formation of an antibody-antigen complex,

(b) imaging any complex formed in this manner,

(c) thereby imaging the presence of activated stromal fibroblasts in a human patient.

62. The method of claim 61, wherein the tumor is a tumor having cancer cells selected from the cancer group consisting of colorectal cancers, non-small cell lung cancers, breast cancers, head and neck cancer, ovarian cancers, lung cancers, bladder cancers, pancreatic cancers and metastatic cancers of the brain.

63. A method of detecting tumor-stroma, characterised in that

- (a) a suitable sample is contacted with an antibody protein according to any one of claims 1 to 24, under conditions suitable for the formation of an antibody-antigen complex,
- (b) detecting the presence of any complex so formed,
- (c) relating the presence of said complex to the presence of tumor-stroma.

5

64. The method of claim 62, wherein said antibody is labelled with a detectable marker.

65. A method of imaging tumor-stroma in a human patient, which comprises

10

- (a) administering to the patient an antibody protein according to any one of claims 42 to 45, under conditions suitable for the formation of an antibody-antigen complex,
- (b) imaging any complex so formed, and thereby imaging the presence of tumor-stroma in a human patient.

15

20

25

30

35

40

45

50

55

Fig. 1

1	11	21	31	41
GACATTGTGA	TGACCCAATC	TCCAGACTCT	TTGGCTGTGT	CTCTAGGGGA
51	61	71	81	91
GAGGGCCACC	ATCAACTGCA	AGTCCAGTCA	GAGCCTTTTA	TATTCTAGAA
101	111	121	131	141
ATCAAAAGAA	CTACTTGGCC	TGGTATCAGC	AGAAACCAGG	ACAGCCACCC
151	161	171	181	191
AAACTCCTCA	TCTTTTGGGC	TAGCACTAGG	GAATCTGGGG	TACCTGATAG
201	211	221	231	241
GTTCAGTGGC	AGTGGGTTTG	GGACAGACTT	CACCCTCACC	ATTAGCAGCC
251	261	271	281	291
TGCAGGCTGA	AGATGTGGCA	GTTTATTACT	GTCAGCAATA	TTTTAGCTAT
301	311	321	331	339
CCGCTCACGT	TCGGACAAGG	GACCAAGGTG	GAAATAAAA	

Fig. 2

1	11	21	31	41
DIVMTQSPDS	LAVSLGERAT	INCKSSQSLL	YSRNQKNYLA	WYQQKPGQPP
51	61	71	81	91
KLLIFWASTR	ESGVPDRFSG	SGFGTDFTLT	ISSLQAEDVA	VYYCQQYFSY
101	111			
PLTFGQGTKV	EIK			

Fig. 3

1	11	21	31	41
GACATTGTGA	TGACCCAATC	TCCAGACTCT	TTGGCTGTGT	CTCTAGGGGA
51	61	71	81	91
GAGGGCCACC	ATCAACTGCA	AGTCCAGTCA	GAGCCTTTTA	TATTCTAGAA
101	111	121	131	141
ATCAAAAGAA	CTACTTGGCC	TGGT <u>T</u> CCAGC	AGAAACCAGG	ACAGCCACCC
151	161	171	181	191
AAACTCCTCA	TCTTTTGGGC	TAGCACTAGG	GAATCTGGGG	TACCTGATAG
201	211	221	231	241
GTTCAGTGGC	AGTGGGTTTG	GGACAGACTT	CACCCTCACC	ATTAGCAGCC
251	261	271	281	291
TGCAGGCTGA	AGATGTGGCA	GTTTAT <u>G</u> ACT	GTCA <u>A</u> CAATA	TTTTAGCTAT
301	311	321	331	339
CCGCTCACGT	TCGGACAAGG	GACCAAGGTG	GAAATAAAA	

Fig. 4

1	11	21	31	41
DIVMTQSPDS	LAVSLGERAT	INCKSSQSL	YSRNQKNYLA	WFQQKPGQPP
51	61	71	81	91
KLLIFWASTR	ESGVPDRFSG	SGFGTDFTLT	ISSLQAEDVA	VYDCQQYFSY
101	111			
PLTFGQGTKV	EIK			

Fig. 5

1	11	21	31	41
GACATTGTGA	TGACCCAATC	TCCAGACTCT	TTGGCTGTGT	CTCTAGGGGA
51	61	71	81	91
GAGGGCCACC	ATCAACTGCA	AGTCCAGTCA	GAGCCTTTTA	TATTCTAGAA
101	111	121	131	141
ATCAAAAGAA	CTACTTGGCC	TGGTATCAGC	AGAAACCAGG	ACAGCCACCC
151	161	171	181	191
AAACTCCTCA	TCTATTGGGC	TAGCACTAGG	GAATCTGGGG	TACCTGATAG
201	211	221	231	241
GTTCAGTGGC	AGTGGGTTTG	GGACAGACTT	CACCCTCACC	ATTAGCAGCC
251	261	271	281	291
TGCAGGCTGA	AGATGTGGCA	GTTTATTACT	GTCAGCAATA	TTTTAGCTAT
301	311	321	331	339
CCGCTCACGT	TCGGACAAGG	GACCAAGGTG	GAAATAAAA	

Fig. 6

1	11	21	31	41
DIVMTQSPDS	LAVSLGERAT	INCKSSQSL	YSRNQKNYLA	WYQQKPGQPP
51	61	71	81	91
KLLIYWASTR	ESGVPDRFSG	SGFGTDFTLT	ISSLQAEDVA	VYYCQQYFSY
101	111			
PLTFGQGTKV	EIK			

Fig. 7

```

1
CAGGTGCAAC TAGTGCAGTC CGGCGCCGAA GTGAAGAAAC CCGGTGCTTC
51
CGTGAAAGTC AGCTGTAAAA CTAGTAGATA CACCTTCACT GAATACACCA
101
TACACTGGGT TAGACAGGCC CCTGGCCAAA GGCTGGAGTG GATAGGAGGT
151
ATTAATCCTA ACAATGGTAT TCCTAACTAC AACCAGAAGT TCAAGGGCCG
201
GGCCACCTTG ACCGTAGGCA AGTCTGCCAG CACCGCCTAC ATGGAAGTGT
251
CCAGCCTGCG CTCCGAGGAC ACTGCAGTCT ACTACTGCGC CAGAAGAAGA
301
ATCGCCTATG GTTACGACGA GGGCCATGCT ATGGACTACT GGGGTCAAGG
351                               372
AACCTTGTC ACCGTCTCCT CA

```

Fig. 8

```

1          11          21          31          41
QVQLVQSGAE VKKPGASVKV SCKTSRYTFT EYTIHWVRQA PGQRLEWIGG
51         61         71         81         91
INPNNGIPNY NQKFKGRATL TVGKSASTAY MELSSLRSED TAVYYCARRR
101        111        121-124
IAYGYDEGHA MDYWGQGLV TVSS

```

Fig. 9

```

1
CAGGTGCAAC TAGTGCAGTC CGGCGCCGAA GTGAAGAAAC CCGGTGCTTC
51
CGTGAAAGTC AGCTGTAAAA CTAGTAGATA CACCTTCACT GAATACACCA
101
TACACTGGGT TAGACAGGCC CCTGGCCAAA GGCTGGAGTG GATAGGAGGT
151
ATTAATCCTA ACAATGGTAT TCCTAACTAC AACCAGAAGT TCAAGGGCCG
201
GGCCACCTTG ACCGTAGGCA AGTCTGCCAG CACCGCCTAC ATGGAAGTGT
251
CCAGCCTGCG CTCCGAGGAC ACTGCAGTCT ACTTCTGCGC CAGAAGAAGA
301
ATCGCCTATG GTTACGACGA GGGCCATGCT ATGGACTACT GGGGTCAAGG
351                               372
AACCTTGTC ACCGTCTCCT CA

```

Fig. 10

1	11	21	31	41
QVQLVQSGAE	VKKPGASVKV	SCKTSRYTFT	EYTIHWVRQA	PGQRLEWIGG
51	61	71	81	91
INPNNGIPNY	NQKFKGRATL	TVGKSASTAY	MELSSLRSED	TAVY <u>F</u> CARRR
101	111	121-124		
IAYGYDEGHA	MDYWGQGTLV	TVSS		

Fig. 11

1				
CAGGTGCAAC	TAGTGCAGTC	CGGCGCCGAA	GTGAAGAAAC	CCGGTGCTTC
51				
CGTGAAAGTC	AGCTGTAAAA	CTAGTAGATA	CACCTTCACT	GAATACACCA
101				
TACACTGGGT	TAGACAGGCC	CCTGGCCAAA	GGCTGGAGTG	GATAGGAGGT
151				
ATTAATCCTA	ACAATGGTAT	TCCTAACTAC	AACCAGAAGT	TCAAGGGCCG
201				
GG <u>T</u> ACCC <u>A</u> <u>T</u> C	ACCGTAG <u>A</u> CA	<u>C</u> CTCTGCCAG	CACCGCCTAC	ATGGAACTGT
251				
CCAGCCTGCG	CTCCGAGGAC	ACTGCAGTCT	ACTACTGCGC	CAGAAGAAGA
301				
ATCGCCTATG	GTTACGACGA	GGGCCATGCT	ATGGACTACT	GGGGTCAAGG
351		372		
AACCCTTGTC	ACCGTCTCCT	CA		

Fig. 12

1	11	21	31	41
QVQLVQSGAE	VKKPGASVKV	SCKTSRYTFT	EYTIHWVRQA	PGQRLEWIGG
51	61	71	81	91
INPNNGIPNY	NQKFKGR <u>V</u> <u>T</u> <u>I</u>	TV <u>D</u> <u>T</u> SASTAY	MELSSLRSED	TAVYYCARRR
101	111	121-124		
IAYGYDEGHA	MDYWGQGTLV	TVSS		

Fig. 13

```

1
CAGGTGCAAC TAGTGCAGTC CGGCGCCGAA GTGAAGAAAC CCGGTGCTTC
51
CGTGAAAGTC AGCTGTAAAA CTAGTAGATA CACCTTCACT GAATACACCA
101
TACACTGGGT TAGACAGGCC CCTGGCCAAA GGCTGGAGTG GATAGGAGGT
151
ATTAATCCTA ACAATGGTAT TCCTAACTAC AACCAGAAGT TCAAGGGCCG
201
GGTCACCAATC ACCGTAGACA CCTCTGCCAG CACCGCCTAC ATGGAAGTGT
251
CCAGCCTGCG CTCCGAGGAC ACTGCAGTCT ACTTCTGCGC CAGAAGAAGA
301
ATCGCCTATG GTTACGACGA GGGCCATGCT ATGGACTACT GGGGTCAAGG
351                               372
AACCCTTGTC ACCGTCTCCT CA

```

Fig. 14

```

1           11           21           31           41
QVQLVQSGAE VKKPGASVKV SCKTSRYTFT EYTIHWVRQA PGQRLEWIGG
51          61          71          81          91
INPNNGIPNY NQKFKGRVTI TVDTSASTAY MELSSLRSED TAVYFCARRR
101         111        121-124
IAYGYDEGHA MDYWGQGTLV TVSS

```

Fig. 15

```

1
CAGGTGCAAC TAGTGCAGTC CGGCGCCGAA GTGAAGAAAC CCGGTGCTTC
51
CGTGAAAGTC AGCTGTAAAA CTAGTGGATA CACCTTCACT GAATACACCA
101
TACACTGGGT TAGACAGGCC CCTGGCCAAA GGCTGGAGTG GATAGGAGGT
151
ATTAATCCTA ACAATGGTAT TCCTAACTAC AACCAGAAGT TCAAGGGCCG
201
GGTCACCAATC ACCGTAGACA CCTCTGCCAG CACCGCCTAC ATGGAAGTGT
251
CCAGCCTGCG CTCCGAGGAC ACTGCAGTCT ACTACTGCGC CAGAAGAAGA
301
ATCGCCTATG GTTACGACGA GGGCCATGCT ATGGACTACT GGGGTCAAGG
351                               372
AACCCTTGTC ACCGTCTCCT CA

```

Fig. 16

1	11	21	31	41
QVQLVQSGAE	VKKPGASVKV	SCKTS <u>G</u> YTFT	EYTIHWVRQA	PGQRLEWIGG
51	61	71	81	91
INPNNGIPNY	NQKF K GR <u>V</u> <u>T</u> <u>I</u>	TV <u>D</u> <u>T</u> SASTAY	MELSSLRSED	TAVYYCARRR
101	111	121-124		
IAYGYDEGHA	MDYWGQGTLV	TVSS		

Fig. 17

1				
DIVMSQSPSS	LAVSVGEKVT	MSCKSSQSLL	YSRNQKNYLA	WFQQKPGQSP
51				
KLLIFWASTR	ESGVPDRFTG	SGFGTDFNLT	ISSVQAEDLA	VYDCQQYFSY
101				
PLTFGAGTKL	ELKRTVAAPS	VFIFPPSDEQ	LKSGTASVVC	LLNNFYPREA
151				
KVQWKVDNAL	QSGNSQESVT	EQDSKDSTYS	LSSTLTLSKA	DYEKHKVYAC
201				
EVTHQGLSSP	VTKSFNRGEC			

Fig. 18

<u>1</u>				
VQLQQSGPEL	VKPGASVKMS	CKTSRYTFTE	YTIHWVRQSH	GKSLEWIGGI
51				
NPNNGIPNYN	QKF K GRATLT	VGKSSSTAYM	ELRSLTSEDS	AVYFCARRRI
101				
AYGYDEGHAM	DYWGQGTSTV	VSSASTKGPS	VFPLAPSSKS	TSGGTAALGC
151				
LVKDYFPEPV	TVSWNSGALT	SGVHTFPAVL	QSSGLYSLSS	VVTVPSSSLG
201				
TQTYICNVNH	KPSNTKVDKK	VEPKSCDKTH	TCPPCPAPEL	LGGPSVFLFP
251				
PKPKDTLMIS	RTPEVTCVVV	DVSHEDPEVK	FNWYVDGVEV	HNAKTKPREE
301				
QYNSTYRVVS	VLTVLHQDWL	NGKEYKCKVS	NKALPAPIEK	TISKAKGQPR
351				
EPQVYTLPPS	REEMTKNQVS	LTCLVKGFYP	SDIAVEWESN	GQPENNYKTT
401				
PPVLDSDGSF	FLYSKLTVDK	SRWQQGNVFS	CSVMHEALHN	HYTQKSLSL
451				
PGK				

Fig. 19

340	350	360	370	380
CGTACTGTGG	CTGCACCATC	TGTCTTCATC	TTCCCGCCAT	CTGATGAGCA
390	400	410	420	430
GTTGAAATCT	GGAAGTGCCT	CTGTTGTGTG	CCTGCTGAAT	AACTTCTATC
440	450	460	470	480
CCAGAGAGGC	CAAAGTACAG	TGGAAGGTGG	ATAACGCCCT	CCAATCGGGT
490	500	510	520	530
AACTCCCAGG	AGAGTGTAC	AGAGCAGGAC	AGCAAGGACA	GCACCTACAG
540	550	560	570	580
CCTCAGCAGC	ACCCTGACGC	TGAGCAAAGC	AGACTACGAG	AAACACAAAG
590	600	610	620	630
TCTACGCCTG	CGAAGTCACC	CATCAGGGCC	TGAGCTCGCC	CGTCACAAAG
640	650	660		
AGCTTCAACA	GGGAGAGTGT			

Fig. 20

114	124	134	144	154
RTVAAPSVFI	FPPSDEQLKS	GTASVVCLLN	NFYPREAKVQ	WKVDNALQSG
164	174	184	194	204
NSQESVTEQD	SKDSTYSLSS	TLTLSKADYE	KHKVYACEVT	HQGLSSPVTK
214-220				
SFNRGEC				

Fig. 21

373
 GCCTCCACCA AGGGCCCATC GGTCTTCCCC CTGGCACCCCT CCTCCAAGAG
 423
 CACCTCTGGG GGCACAGCGG CCCTGGGCTG CCTGGTCAAG GACTACTTCC
 473
 CCGAACCGGT GACGGTGTCG TGGAAGTCAG GCGCCCTGAC CAGCGGCGTG
 523
 CACACCTTCC CGGCTGTCCT ACAGTCCTCA GGACTCTACT CCCTCAGCAG
 573
 CGTGGTGACC GTGCCCTCCA GCAGCTTGGG CACCCAGACC TACATCTGCA
 623
 ACGTGAATCA CAAGCCCAGC AACACCAAGG TGGACAAGAA AGTTGAGCCC
 673
 AAATCTTGTG ACAAAGTCAC CACATGCCCA CCGTGCCCAG CACCTGAACT
 723
 CCTGGGGGGA CCGTCAGTCT TCCTCTTCCC CCCAAAACCC AAGGACACCC
 773
 TCATGATCTC CCGGACCCCT GAGGTCACAT GCGTGGTGGT GGACGTGAGC
 823
 CACGAAGACC CTGAGGTCAA GTTCAACTGG TACGTGGACG GCGTGGAGGT
 873
 GCATAATGCC AAGACAAAGC CGCGGGAGGA GCAGTACAAC AGCACGTACC
 923
 GGGTGGTCAG CGTCCTCACC GTCCTGCACC AGGACTGGCT GAATGGCAAG
 973
 GAGTACAAGT GCAAGGTCTC CAACAAAGCC CTCCCAGCCC CCATCGAGAA
 1023
 AACCATCTCC AAAGCCAAAG GGCAGCCCCG AGAACCACAG GTGTACACCC
 1073
 TGCCCCCATC CCGGGAGGAG ATGACCAAGA ACCAGGTCAG CCTGACCTGC
 1123
 CTGGTCAAAG GCTTCTATCC CAGCGACATC GCCGTGGAGT GGGAGAGCAA
 1173
 TGGGCAGCCG GAGAACAAC TACAAGACCAC GCCTCCCGTG CTGGACTCCG
 1223
 ACGGCTCCTT CTCCTCTAC AGCAAGCTCA CCGTGGACAA GAGCAGGTGG
 1273
 CAGCAGGGGA ACGTCTTCTC ATGCTCCGTG ATGCATGAGG CTCTGCACAA
 1323
 CCACTACACG CAGAAGAGCC TCTCCCTGTC TCCGGGTAAA
 1362

Fig. 22

125
 ASTKGPSVFP LAPSSKSTSG GTAALGCLVK DYFPEPVTVS WNSGALTSGV
 175
 HTFPAVLQSS GLYSLSSVVT VPSSSLGTQT YICNVN HKPS NTKVDKKVEP
 225
 KSCDKTHTCP PCPAPELLGG PSVFLFPPKP KDTLMIS RTP EVTCVVVDVS
 275
 HEDPEVKFNW YVDGVEVHNA KTKPREEQYN STYRVVSVLT VLHQDWLNGK
 325
 EYKCKVSNKA LPAPIEKTIS KAKGQPREPQ VYTLPPSREE MTKNQVSLTC
 375
 LVKGFYPSDI AVEWESNGQP ENNYKTTPPV LDSDGSFFLY SKLTVDKSRW
 425 454
 QQGNVFSCSV MHEALHNHYT QKSLSLSPGK

Fig. 23A

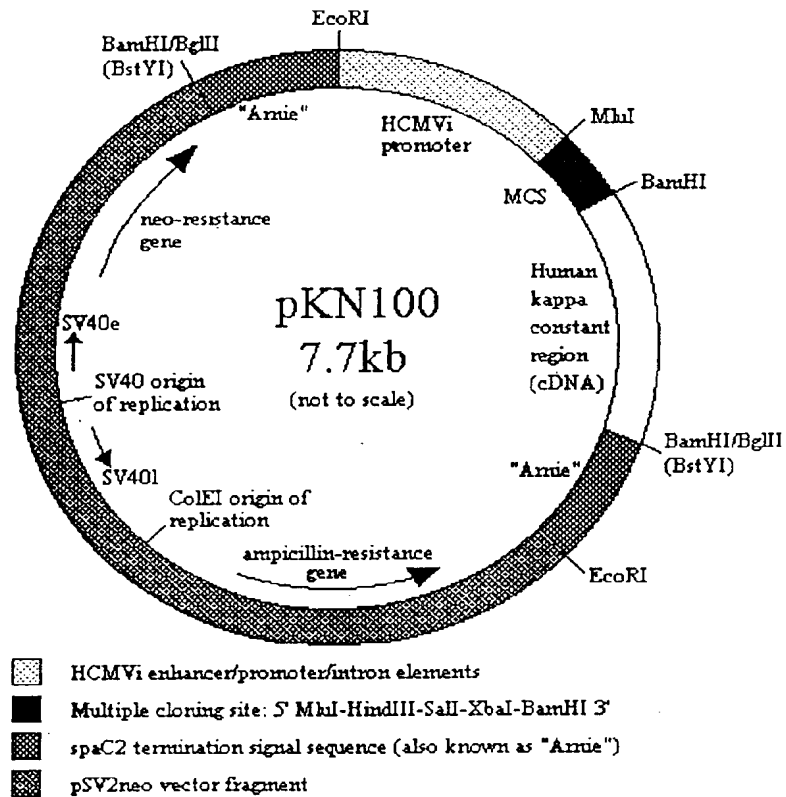


Fig. 23B

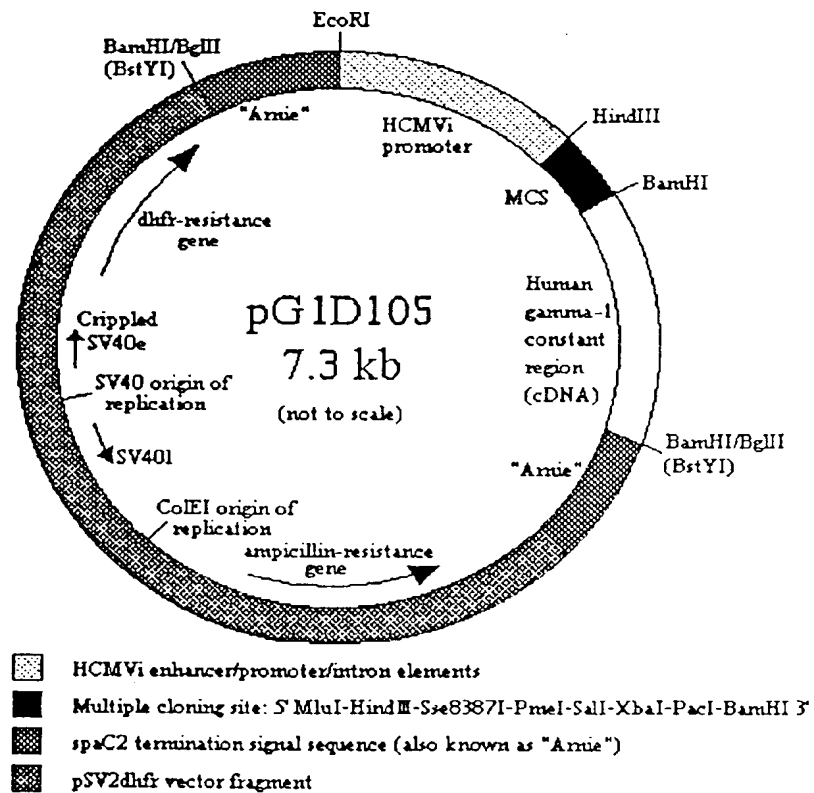


Fig. 24

HindIII
aagcttGCCGCCACCatggattcacaggcccaggttctttatgttactgccgctatgggta
 1 -----+-----+-----+-----+-----+-----+
ttcgaaCGGCGGTGGtacctaagtgtccgggtccaagaataacaatgacggcgatacccat
 Kozak sequence
 M D S Q A Q V L M L L P L W V

tctggtagctgtggggacattgtgatgtcacagtctccatcctccctagctgtgtcagtt
 61 -----+-----+-----+-----+-----+-----+
 agaccatggacacccctgtaacactacagtgtcagaggtaggaggatcgacacagtcaa
 S G T C G D I V M S Q S P S S L A V S V

ggagagaagggttactatgagctgcaagtccagtcagagccttttatatagtcgtaataca
 121 -----+-----+-----+-----+-----+-----+
 cctctcttccaatgatactcgacgttcaggtcagtcctcggaataatatacagcattagtt
 G E K V T M S C K S S Q S L L Y S R N Q
 CDR 1

aagaactacttggcctggttccagcagaagccagggcagtcctcctaaactgctgattttc
 181 -----+-----+-----+-----+-----+-----+
 ttcttgatgaaccggaccaaggtcgtcttcggtcccgtcagaggatttgacgactaaaag
 K N Y L A W F Q Q K P G Q S P K L L I F

tgggcatccactaggggaatctgggggtccctgatcgcttcacaggcagtggtttgggacg
 241 -----+-----+-----+-----+-----+-----+
 acccgtaggtgatcccttagacccccagggactagcgaagtgtccgtcacctaaaccctgc
 W A S T R E S G V P D R F T G S G F G T
 CDR 2

gatttcaatctcaccatcagcagtggtgcaggctgaggacctggcagtttatgactgtcag
 301 -----+-----+-----+-----+-----+-----+
 ctaaagtttagagtggtagtcgtcacacgtccgactcctggaccgtcaaatactgacagtc
 D F N L T I S S V Q A E D L A V Y D C Q

caatatttttagctatccgctcacgttcggtgctgggaccaagctggagctgaAACGTGAG
 361 -----+-----+-----+-----+-----+-----+
 gttataaaatcgataggcgagtgcaagccacgaccctgggttcgacctcgactTTGCACTG
 splice donor site
 Q Y F S Y P L T F G A G T K L E L K
 CDR 3

BamHI
Tggatcc
 421 ----- 427
Acctagg

Fig. 25

HindIII
AAGCTTGCCGCCACCATGGGATGGAGCTGGGTCTTTCTCTTTCTCCTGTCAGGAACTGCA
 1 -----+-----+-----+-----+-----+-----+-----+
TCGAAACGGCGGTGGTACCCTACCTCGACCCAGAAAGAGAAAGAGGACAGTCCTTGACGT
Kozak sequence
 M G W S W V F L F L L S G T A

GGTGTCTCTCTGAGGTCCAGCTGCAACAGTCTGGACCTGAGCTGGTGAAGCCTGGGGCT
 61 -----+-----+-----+-----+-----+-----+-----+
 CCACAGGAGAGACTCCAGGTTCGACGTTGTCAGACCTGGACTCGACCACTTCGGACCCCGA
 G V L S E V Q L Q Q S G P E L V K P G A

TCAGTAAAGATGTCCTGCAAGACTTCTAGATACACATTCAGTGAATACACCATACACTGG
 121 -----+-----+-----+-----+-----+-----+-----+
 AGTCATTTCTACAGGACGTTCTGAAGATCTATGTGTAAGTGACTTATGTGGTATGTGACC
 S V K M S C K T S R Y T F T E Y T I H W
 CDR 1

GTGAGACAGAGCCATGGAAAGAGCCTTGAGTGGATTGGAGGTATTAATCCTAACAATGGT
 181 -----+-----+-----+-----+-----+-----+-----+
 CACTCTGTCTCGGTACCTTTCTCGGAACTCACCTAACCTCCATAATTAGGATTGTTACCA
 V R Q S H G K S L E W I G G I N P N N G
 CDR 2

ATTCCTAACTACAACCAGAAGTTCAAGGGCAGGGCCACATTGACTGTAGGCAAGTCCTCC
 241 -----+-----+-----+-----+-----+-----+-----+
 TAAGGATTGATGTTAGTCTTCAAGTTCCCGTCCCGGTGTAAGTGACATCCGTTCCAGGAGG
 I P N Y N Q K F K G R A T L T V G K S S

AGCACCGCCTACATGGAGCTCCGCAGCCTGACATCTGAGGATTCTGCGGTCTATTTCTGT
 301 -----+-----+-----+-----+-----+-----+-----+
 TCGTGGCGGATGTACCTCGAGGCGTCGGACTGTAGACTCCTAAGACGCCAGATAAAGACA
 S T A Y M E L R S L T S E D S A V Y F C

GCAAGAAGAAGAATCGCCTATGGTTACGACGAGGGCCATGCTATGGACTACTGGGGTCAA
 361 -----+-----+-----+-----+-----+-----+-----+
 CGTTCTTCTTCTTAGCGGATACCAATGCTGCTCCCGGTACGATACCTGATGACCCAGTT
 A R R R I A Y G Y D E G H A M D Y W G Q
 CDR 3

BamHI
 GGAACCTCAGTCACCGTCTCCTCAGGTGAGT**GGATCC**
 421 -----+-----+-----+-----+-----+-----+-----+ 468
 CCTTGGAGTCAGTGGCAGAGGAGTCCACTCACCTAGG
 splice donor site
 G T S V T V S S

Fig. 26 /1

```

      Spe I
1   gaattccagc acactggcgg ccgttACTAG TTATTAATAG TAATCAATTA
51  CGGGGTCATT AGTTCATAGC CCATATATGG AGTTCCGCGT TACATAACTT
101 ACGGTAAATG GCCCGCCTGG CTGACCGCCC AACGACCCCC GCCCATTGAC
151 GTCAATAATG ACGTATGTTC CCATAGTAAC GCCAATAGGG ACTTTCCATT
201 GACGTCAATG GGTGGAGTAT TTACGGTAAA CTGCCCCTT GGCAGTACAT
251 CAAGTGTATC ATATGCCAAG TACGCCCCCT ATTGACGTCA ATGACGGTAA
301 ATGGCCCCGCC TGGCATTATG CCCAGTACAT GACCTTATGG GACTTTCCTA
      SnaB I
351 CTTGGCAGTA CATCTACGTA TTAGTCATCG CTATTACCAT GGTGATGCGG
401 TTTTGGCAGT ACATCAATGG GCGTGGATAG CGGTTTGACT CACGGGGATT
451 TCCAAGTCTC CACCCCATTTG ACGTCAATGG GAGTTTGTTT TGGCACCAAA
501 ATCAACGGGA CTTTCCAAAA TGTCGTAACA ACTCCGCCCC ATTGACGCAA
551 ATGGGCGGTA GGCGTGTACG GTGGGAGGTC TATATAAGCA GAGCTCGTTT
601 AGTGAACCGT CAGATCGCCT GGAGACGCCA TCCACGCTGT TTTGACCTCC
      Sac II
651 ATAGAAGACA CCGGGACCGA TCCAGCCTCC GCGGCCGGGA ACGGTGCATT
701 GGAACGCGGA TTCCCCGTGC CAAGAGTGAC GTAAGTACCG CCTATAGAGT
751 CTATAGGCCC ACCCCCTTGG CTTCTTATGC ATGCTATACT GTTTTTGGCT
801 TGGGGTCTAT ACACCCCCGC TTCCTCATGT TATAGGTGAT GGTATAGCTT
851 AGCCTATAGG TGTGGGTTAT TGACCATTAT TGACCACTCC CCTATTGGTG
901 ACGATACTTT CCATTACTAA TCCATAACAT GGCTCTTTGC CACAACTCTC
951 TTTATTGGCT ATATGCCAAT AACTGTCCT TCAGAGACTG ACACGGACTC
1001 TGTATTTTTA CAGGATGGGG TCTCATTTAT TATTTACAAA TTCACATATA
1051 CAACACCACC GTCCCCAGTG CCCGCAGTTT TTATTAAACA TAACGTGGGA
      BspE I
1101 TCTCCACGCG AATCTCGGGT ACGTGTTCCG GACATGGGCT CTTCTCCGGT
1151 AGCGGCGGAG CTTCTACATC CGAGCCCTGC TCCCATGCCT CCAGCGACTC
1201 ATGGTCGCTC GGCAGCTCCT TGCTCCTAAC AGTGGAGGCC AGACTTAGGC
1251 ACAGCACGAT GCCCACCACC ACCAGTGTGC CGCACAAGGC CGTGGCGGTA

```

Fig. 26 /2

1301	GGGTATGTGT	CTGAAAATGA	GCTCgggggag	cgggcttgca	ccgctgacgc	
		Afl II				
1351	atttggaaga	<u>cttaag</u> gcag	cggcagaaga	agatgcaggc	agctgagttg	
1401	ttgtgttctg	ataagagtca	gaggtaactc	ccgttgcggt	gctgttaacg	
1451	gtggagggca	gtgtagtctg	agcagtactc	gttgctgccg	cgcgcgccac	
1501	cagacataat	agctgacaga	ctaacagact	gttcctttcc	atgggtcttt	
			Mlu I	Hind III		
1551	tctgcagtca	ccgtccttga	<u>cacgcg</u> tctc	ggga <u>aagctt</u> G	CCGCCACCAT	
					M	
					Kpn I	
1601	GGATTCACAG	GCCCAGGTTC	TTATGTTACT	GCCGCTATGG	GTATCT <u>GGTA</u>	
	D S Q	A Q V	L M L L	P L W	V S G	
1651	<u>CCT</u> GTGGGGA	CATTGTGATG	TCACAGTCTC	CATCCTCCCT	AGCTGTGTCA	
	T C G D	I V M	S Q S	P S S L	A V S	
1701	GTTGGAGAGA	AGGTTACTAT	GAGCTGCAAG	TCCAGTCAGA	GCCTTTTATA	
	V G E	K V T M	S C K	S S Q	S L L Y	
	XbaI			CDR 1		
1751	<u>TTCTAGAA</u> AT	CAAAAGAACT	ACTTGGCCTG	GTTCCAGCAG	AAGCCAGGGC	
	S R N	Q K N	Y L A W	F Q Q	K P G	
1801	AGTCTCCTAA	ACTGCTGATT	TTCTGGGCAT	CCACTAGGGA	ATCTGGGGTC	
	Q S P K	L L I	F W A	S T R E	S G V	
				CDR 2		
1851	CCTGATCGCT	TCACAGGCAG	TGGATTGTTGGG	ACGGATTTC	ATCTCACC	
	P D R	F T G S	G F G	T D F	N L T I	
1901	CAGCAGTGTG	CAGGCTGAGG	ACCTGGCAGT	TTATGACTGT	CAGCAATATT	
	S S V	Q A E	D L A V	Y D C	<u>Q Q Y</u>	
1951	TTAGCTATCC	GCTCACGTTTC	GGTGCTGGGA	CCAAGCTGGA	GCTGAAACGT	
	<u>F S Y P</u>	<u>L T F</u>	G A G	T K L E	L K R	
	CDR 3					
	BamH I					
2001	<u>GAGTggatcc</u>	ATCTGGGATA	AGCATGCTGT	TTTCTGTCTG	TCCCTAACAT	
2051	GCCCTGTGAT	TATGCGCAAA	CAACACACCC	AAGGGCAGAA	CTTTGTTACT	
2101	TAAACACCAT	CCTGTTTGCT	TCTTTCCTCA	<u>GGA</u> ACTGTGG	CTGCACCATC	
				T V	A A P S	
2151	TGTCTTCATC	TTCCCGCCAT	CTGATGAGCA	GTTGAAATCT	GGAAGTGCCT	
	V F I	F P P	S D E Q	L K S	G T A	
2201	CTGTTGTGTG	CCTGCTGAAT	AACTTCTATC	CCAGAGAGGC	CAAAGTACAG	
	S V V C	L L N	N F Y	P R E A	K V Q	
2251	TGGAAGGTGG	ATAACGCCCT	CCAATCGGGT	AACTCCCAGG	AGAGTGTGAC	
	W K V	D N A L	Q S G	N S Q	E S V T	
2301	AGAGCAGGAC	AGCAAGGACA	GCACCTACAG	CCTCAGCAGC	ACCCTGACGC	
	E Q D	S K D	S T Y S	L S S	T L T	

Fig. 26 /3

```

2351  TGAGCAAAGC AGACTACGAG AAACACAAAG TCTACGCCTG CGAAGTCACC
      L S K A   D Y E   K H K   V Y A C   E V T
2401  CATCAGGGCC TGAGCTCGCC CGTCACAAAG AGCTTCAACA GGGGAGAGTG
      H Q G   L S S P   V T K   S F N   R G E C
2451  TTAGAGGGAG AAGTGCCCCC ACCTGCTCCT CAGTTCCAGC CTGACCCCCT
      *
2501  CCCATCCTTT GGCCTCTGAC CCTTTTTTCCA CAGGGGACCT ACCCCTATTG
2551  CGGTCCTCCA GCTCATCTTT CACCTCACCC CCCTCCTCCT CCTTGGCTTT
2601  AATTATGCTA ATGTTGGAGG AGAATGAATA AATAAAGTGA ATCTTTGCAC
2651  CTGTGGTGGA TCTAATAAAA GATATTTATT TTCATTAGAT ATGTGTGTTG
2701  GTTTTTTGTG TGCAGTGCCT CTATCTGGAG GCCAGGTAGG GCTGGCCTTG
2751  GGGGAGGGGG AGGCCAGAAT GACTCCAAGA GCTACAGGAA GGCAGGTCAG
2801  AGACCCCACT GGACAAACAG TGGCTGGACT CTGCACCATA ACACACAATC
2851  AACAGGGGAG TGAGCTGGAA ATTTGCTAGC GAATTCTTGA AGACGAAAGG
2901  GCCTCGTGAT ACGCCTATTT TTATAGGTTA ATGTCATGAT AATAATGGTT
2951  TCTTAGACGT CAGGTGGCAC TTTTCGGGGA AATGTGCGCG GAACCCCTAT
3001  TTGTTTATTT TTCTAAATAC ATTCAAATAT GTATCCGCTC ATGAGACAAT
3051  AACCTTGATA AATGCTTCAA TAATATTGAA AAAGGAAGAG TATGAGTATT
3101  CAACATTTCC GTGTCGCCCT TATTCCCTTT TTTGCGGCAT TTTGCCTTCC
3151  TGTTTTTTGCT CACCCAGAAA CGCTGGTGAA AGTAAAAGAT GCTGAAGATC
3201  AGTTGGGTGC ACGAGTGGGT TACATCGAAC TGGATCTCAA CAGCGGTAAG
3251  ATCCTTGAGA GTTTTCGCCC CGAAGAACGT TTTCCAATGA TGAGCACTTT
3301  TAAAGTTCTG CTATGTGGCG CGGTATTATC CCGTGTTGAC GCCGGGCAAG
3351  AGCAACTCGG TCGCCGCATA CACTATTCTC AGAATGACTT GGTTGAGTAC
3401  TCACCAGTCA CAGAAAAGCA TCTTACGGAT GGCATGACAG TAAGAGAATT
3451  ATGCAGTGCT GCCATAACCA TGAGTGATAA CACTGCGGCC AACTTACTTC
      Pvu I
3501  TGACAACCGAT CGGAGGACCG AAGGAGCTAA CCGCTTTTTT GCACAACATG
3551  GGGGATCATG TAACTCGCCT TGATCGTTGG GAACCGGAGC TGAATGAAGC
3601  CATACCAAAC GACGAGCGTG ACACCACGAT GCCTGCAGCA ATGGCAACAA

```


Fig. 26 /4

3651 CGTTGCGCAA ACTATTAACT GGCGAACTAC TTA CTCTAGC TTCCCGGCAA
 3701 CAATTAATAG ACTGGATGGA GGCGGATAAA GTTGCAGGAC CACTTCTGCG
 3751 CTCGGCCCTT CCGGCTGGCT GGT TTATTGC TGATAAATCT GGAGCCGGTG
 3801 AGCGTGGGTC TCGCGGTATC ATTGCAGCAC TGGGGCCAGA TGGTAAGCCC
 3851 TCCCGTATCG TAGTTATCTA CACGACGGGG AGTCAGGCAA CTATGGATGA
 3901 ACGAAATAGA CAGATCGCTG AGATAGGTGC CTCACTGATT AAGCATTTGGT
 3951 AACTGTCAGA CCAAGTTTAC TCATATATAC TTTAGATTGA TTTAAAACTT
 4001 CATTTTTAAT TTAAAAGGAT CTAGGTGAAG ATCCTTTTTG ATAATCTCAT
 4051 GACCAAATC CCTTAACGTG AGTTTTTCGTT CCACTGAGCG TCAGACCCCG
 4101 TAGAAAAGAT CAAAGGATCT TCTTGAGATC CTTTTTTTCT GCGCGTAATC
 4151 TGCTGCTTGC AAACAAAAAA ACCACCGCTA CCAGCGGTGG TTTGTTTGCC
 4201 GGATCAAGAG CTACCAACTC TTTTTCGAA GGTA ACTGGC TTCAGCAGAG
 4251 CGCAGATACC AAATACTGTC CTTCTAGTGT AGCCGTAGTT AGGCCACCAC
 4301 TTCAAGAACT CTGTAGCACC GCCTACATAC CTCGCTCTGC TAATCCTGTT
 4351 ACCAGTGGCT GCTGCCAGTG GCGATAAGTC GTGTCTTACC GGGTTGGACT
 4401 CAAGACGATA GTTACCGGAT AAGGCGCAGC GGTCGGGCTG AACGGGGGGT
 4451 TCGTGACAC AGCCAGCTT GGAGCGAACG ACCTACACCG AACTGAGATA
 4501 CCTACAGCGT GAGCTATGAG AAAGCGCCAC GCTTCCCGAA GGGAGAAAGG
 4551 CGGACAGGTA TCCGGTAAGC GGCAGGGTCG GAACAGGAGA GCGCACGAGG
 4601 GAGCTTCCAG GGGGAAACGC CTGGTATCTT TATAGTCCTG TCGGGTTTCG
 4651 CCACCTCTGA CTTGAGCGTC GATTTTTGTG ATGCTCGTCA GGGGGGCGGA
 4701 GCCTATGGAA AAACGCCAGC AACGCGCCT TTTTACGGTT CCTGGCCTTT
 4751 TGCTGGCCTT TTGCTCACAT GTTCTTTTCCT GCGTTATCCC CTGATTCTGT
 4801 GGATAACCGT ATTACCGCCT TTGAGTGAGC TGATACCGCT CGCCGCAGCC
 4851 GAACGACCGA GCGCAGCGAG TCAGTGAGCG AGGAAGCGGA AGAGCGCCTG
 4901 ATGCGGTATT TTCTCCTTAC GCATCTGTGC GGTATTTTAC ACCGCATATG

Fig. 26 /5

4951 GTGCACTCTC AGTACAATCT GCTCTGATGC CGCATAGTTA AGCCAGTATA ^{Bst1107I}
 5001 CACTCCGCTA TCGCTACGTG ACTGGGTCAT GGCTGCGCCC CGACACCCGC
 5051 CAACACCCGC TGACGCGCCC TGACGGGCTT GTCTGCTCCC GGCATCCGCT
 5101 TACAGACAAG CTGTGACCGT CTCCGGGAGC TGCATGTGTC AGAGGTTTTTC
 5151 ACCGTCATCA CCGAAACGCG CGAGGCAGCT GTGGAATGTG TGTCAGTTAG
 5201 GGTGTGGAAA GTCCCCAGGC TCCCCAGCAG GCAGAAGTAT GCAAAGCATG
 5251 CATCTCAATT AGTCAGCAAC CAGGCTCCCC AGCAGGCAGA AGTATGCAAA
 5301 GCATGCATCT CAATTAGTCA GCAACCATAG TCCCGCCCCCT AACTCCGCCC
 5351 ATCCCGCCCC TAACTCCGCC CAGTTCCGCC CATTCTCCGC CCCATGGCTG
 5401 ACTAATTTTT TTTATTTATG CAGAGGCCGA ^{Sfi I} GGCCGCCTCG GCCTCTGAGC
 5451 TATTCCAGAA GTAGTGAGGA GGCTTTTTTG ^{Stu I/Avr II} GAGGCCTAGG CTTTTGCAAA
 5501 AAGCTAGCTT CACGCTGCCG CAAGCACTCA GGGCGCAAGG GCTGCTAAAG
 5551 GAAGCGGAAC ACGTAGAAAG CCAGTCCGCA GAAACGGTGC TGACCCCGGA
 5601 TGAATGTCAG CTACTGGGCT ATCTGGACAA GGGAAAACGC AAGCGCAAAG
 5651 AGAAAGCAGG TAGCTTGCAAG TGGGCTTACA TGGCGATAGC TAGACTGGGC
 5701 GGTTTTATGG ACAGCAAGCG AACCAGGAATT GCCAGCTGGG GCGCCCTCTG
 5751 GTAAGGTTGG GAAGCCCTGC AAAGTAAACT GGATGGCTTT CTTGCCGCCA
 5801 AGGATCTGAT GCGCGAGGGG ^{Bgl II/Bcl I} ATCAAGATCT GATCAAGAGA CAGGATGAGG
 5851 ATCGTTTCGC ATGATTGAAC AAGATGGATT GCACGCAGGT TCTCCGGCCG
 5901 CTTGGGTGGA GAGGCTATTC GGCTATGACT GGGCACAACA GACAATCGGC
 5951 TGCTCTGATG CCGCCGTGTT CCGGCTGTCA GCGCAGGGGC GCCCGGTTCT
 6001 TTTTGTCAAG ACCGACCTGT CCGGTGCCCT GAATGAACTG CAGGACGAGG
 6051 CAGCGCGGCT ATCGTGGCTG ^{Msc I} GCCACGACGG GCGTTCCTTG CGCAGCTGTG
 6101 CTCGACGTTG TCACTGAAGC GGGAAGGGAC TGGCTGCTAT TGGGCGAAGT
 6151 GCCGGGGCAG GATCTCCTGT CATCTCACCT TGCTCCTGCC GAGAAAGTAT
 6201 CCATCATGGC TGATGCAATG CGGCGGCTGC ATACGCTTGA TCCGGCTACC

Fig. 26 /6

6251 TGCCCATTCG ACCACCAAGC GAAACATCGC ATCGAGCGAG CACGTACTCG
 6301 GATGGAAGCC GGTCTTGTCG ATCAGGATGA TCTGGACGAA GAGCATCAGG
 6351 GGCTCGCGCC AGCCGAACTG TTCGCCAGGC TCAAGGCGCG CATGCCCCGAC
 6401 GGCGAGGATC TCGTCGTGAC CCATGGCGAT GCCTGCTTGC CGAATATCAT
 6451 GGTGGA^{Rsr II}AAAT GGCCGCTTTT CTGGATTTCAT CGACTGTGGC CGGCTGGGTG
 6501 TGGCGGACCG CTATCAGGAC ATAGCGTTGG CTACCCGTGA TATTGCTGAA
 6551 GAGCTTGGCG GCGAATGGGC TGACCGCTTC CTCGTGCTTT ACGGTATCGC
 6601 CGCTCCCGAT TCGCAGCGCA TCGCCTTCTA TCGCCTTCTT GACGAGTTCT
 6651 TCTGAGCGGG ACTCTGGGGT ^{Nsp V}TCGAAATGAC CGACCAAGCG ACGCCCAACC
 6701 TGCCATCACG AGATTTTCGAT TCCACCGCCG CCTTCTATGA AAGGTGGGGC
 6751 TTCGGAATCG TTTTCCGGGA CGCCGGCTGG ATGATCCTCC AGCGCGGGGA
 6801 TCTCATGCTG GAGTTCTTCG ^{Sma I}CCCACCCCCG ^{Nru I}GCTCGATCCC CTCGCGAGTT
 6851 GGTTCACTG CTGCCTGAGG CTGGACGACC TCGCGGAGTT CTACCGGCAG
 6901 TGCAAATCCG TCGGCATCCA GGAAACCAGC AGCGGCTATC CGCGCATCCA
 6951 TGCCCCCGAA CTGCAGGAGT GGGGAGGCAC GATGGCCGCT TTGGTCCCGG
 7001 ATCTTTGTGA AGGAACCTTA CTTCTGTGGT GTGACATAAT TGGACAAACT
 7051 ACCTACAGAG ATTTAAAGCT CTAAGGTAAA TATAAAATTT TTAAGTGTAT
 7101 AATGTGTTAA ACTACTGATT CTAATTGTTT GTGTATTTTA GATTCCAACC
 7151 TATGGA^{Rsr II}ACTG ATGAATGGGA GCAGTGGTGG AATGCCTTTA ATGAGGAAAA
 7201 CCTGTTTTGC TCAGAAGAAA TGCCATCTAG TGATGATGAG GCTACTGCTG
 7251 ACTCTCAACA TTCTACTCCT CCAAAAAAGA AGAGAAAGGT AGAAGACCCC
 7301 AAGGACTTTC CTTCAGAATT GCTAAGTTTT TTGAGTCATG CTGTGTTTAG
 7351 TAATAGA^{Rsr II}ACT CTTGCTTGCT TTGCTATTTA CACCACAAAG GAAAAAGCTG
 7401 CACTGCTATA CAAGAAAATT ATGGAAAAAT ATTCTGTAAC CTTTATAAGT
 7451 AGGCATAACA GTTATAATCA TAACATACTG TTTTTTCTTA CTCCACACAG
 7501 GCATAGAGTG TCTGCTATTA ATA^{Rsr II}ACTATGC TCAAAAATTG TGTACCTTTA

Fig. 26 /7

7551 GCTTTTAAAT TTGTAAAGGG GTTAATAAGG AATATTGAT GTATAGTGCC
 7601 TTGACTAGAG ATCATAATCA GCCATACCAC ATTTGTAGAG GTTTTACTTG
 7651 CTTTAAAAAA CCTCCACAC CTCCCCCTGA ACCTGAAACA TAAAATGAAT
 Mun I
 7701 GCAATTGTTG TTGTAACTT GTTTATTGCA GCTTATAATG GTTACAAATA
 7751 AAGCAATAGC ATCACAAATT TCACAAATAA AGCATTTTTT TCACTGCATT
 7801 CTAGTTGTGG TTTGTCCAAA CTCATCAATG TATCTTATCA TGTCTGGATC
 7851 TAATAAAAGA TATTTATTTT CATTAGATAT GTGTGTTGGT TTTTGTGTG
 7901 CAGTGCCTCT ATCTGGAGGC CAGGTAGGGC TGGCCTTGGG GGAGGGGGAG
 7951 GCCAGAATGA CTCCAAGAGC TACAGGAAGG CAGGTCAGAG ACCCCTACTGG
 8001 ACAAACAGTG GCTGGACTCT GCACCATAAC ACACAATCAA CAGGGGAGTG
 8051 AGCTGGAAAT TTGCTAGC

Fig. 27/1

1 TTGAAGACGAAAGGGCCTCGTGATACGCCTATTTTTATAGGTTAATGTCATGATAATAAT
 61 GGTTCCTTAGACGTCAGGTGGCACTTTTCGGGGAAATGTGCGCGGAACCCCTATTTGTTT
 121 ATTTTTCTAAATACATTCAAATATGTATCCGCTCATGAGACAATAACCCCTGATAAATGCT
 181 TCAATAATATTGAAAAAGGAAGAGTATGAGTATTCAACATTTCCGTGTCGCCCTTATTCC
 241 CTTTTTTCGCGCATTTTGCCTTCCTGTTTTTGCTCACCCAGAAACGCTGGTGAAAGTAAA
 301 AGATGCTGAAGATCAGTTGGGTGCACGAGTGGGTACATCGAACTGGATCTCAACAGCGG
 361 TAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTCCAATGATGAGCACTTTTAAAGT
 421 TCTGCTATGTGGCGCGGTATTATCCCGTGTTGACGCCGGCAAGAGCAACTCGGTGCGCG
 481 CATACACTATTCTCAGAATGACTTGGTTGAGTACTCACCAGTCACAGAAAAGCATCTTAC
 541 GGATGGCATGACAGTAAGAGAATTATGCAGTGCTGCCATAACCATGAGTGATAAACTGC
 Pvu I
 601 GGCCAACTTACTTCTGACAACGATCGGGAGGACCGAAGGAGCTAACCGCTTTTTTGACAA
 661 CATGGGGGATCATGTAACTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACC
 Fsp I
 721 AAACGACGAGCGTGACACCACGATGCCTGCAGCAATGGCAACAACGTTGCGCAAACTATT
 781 AACTGGCGAACTACTTACTCTAGCTTCCCGCAACAATTAATAGACTGGATGGAGGCGGA

Fig. 27 /2

841 TAAAGTTGCAGGACCACTTCTGCGCTCGGCCCTTCCGGCTGGCTGGTTTATTGCTGATAA
 901 ATCTGGAGCCGGTGAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAA
 961 GCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGGATGAACGAAA
 1021 TAGACAGATCGCTGAGATAGGTGCCTCACTGATTAAGCATTTGGTAACTGTCAGACCAAGT
 1081 TTAATCATATATACTTTAGATTGATTTAAAACTTCATTTTAAATTTAAAAGGATCTAGGT
 1141 GAAGATCCTTTTTTGATAATCTCATGACCAAAATCCCTTAACGTGAGTTTTCTGTTCCACTG
 1201 AGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTTCTGCGCGT
 1261 AATCTGCTGCTTGCAAACAAAAAAACCACCGCTACCAGCGGTGGTTTGTGGCCGGATCA
 1321 AGAGCTACCAACTCTTTTTCCGAAGGTAAGTGGCTTCAGCAGAGCGCAGATACCAAATAC
 1381 TGTCTTCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTAC
 1441 ATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCT
 1501 TACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTGGGCTGAACGGG
 1561 GGGTTCTGTCACACAGCCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCTACA
 1621 GCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGT
 1681 AAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGGTA
 1741 TCTTTATAGTCCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTC
 1801 GTCAGGGGGGCGGAGCCTATGGAAAAACGCCAGCAACGCGGCCTTTTTACGGTTCCTGGC
 1861 CTTTTGCTGGCCTTTTGCTCACATGTTCCTTCCTGCGTTATCCCCTGATTCTGTGGATAA
 1921 CCGTATTACCGCCTTTGAGTGAGCTGATACCGCTCGCCGCAGCCGAACGACCGAGCGCAG
 1981 CGAGTCAGTGAGCGAGGAAGCGGAAGAGCGCCTGATGCGGTATTTCTCCTTACGCATCT
 2041 GTGCGGTATTTACACCGCATATGGTGCACCTCTCAGTACAATCTGCTCTGATGCCGCATA
 2101 GTTAAGCCAGGTATACACTCCGCTATCGCTACGTGACTGGGTCATGGCTGCGCCCCGACAC
 2161 CCGCCAACACCCGCTGACGCGCCCTGACGGGCTTGTCTGCTCCCGGCATCCGCTTACAGA
 2221 CAAGCTGTGACCGTCTCCGGGAGCTGCATGTGTGAGAGGTTTTACCGTCATCACCGAAA
 2281 CGCGCGAGGCAGCATGCATCTCAATTAGTCAGCAACCATAGTCCCGCCCCTAACCTCCGCC
 2341 CATCCCGCCCCTAACCTCCGCCAGTTCCGCCCATTTCTCCGCCCATGGCTGACTAATTTT
 2401 TTTTATTTATGCAGAGGCCGAGGGCCGCCTCGGCCTCTGAGCTATTCCAGAAGTAGTGAGG
 2461 AGGCTTTTTTGGAGGCCTAGGCTTTTGCAAAAAGCTAGCTTACAGCTCAGGGCTGCGATT

Fig. 27 /3

2521 TCGCGCCAAACTTGACGGCAATCCTAGCGTGAAGGCTGGTAGGATTTTATCCCCGCTGCC
 2581 ATCATGGTTCGACCATTGAACTGCATCGTCGCCGTGTCCCAAATATGGGGATTGGCAAG
 2641 AACGGAGACCTACCCTGGCCTCCGCTCAGGAACGAGTTCAAGTACTTCCAAAGAATGACC
 2701 ACAACCTCTTCAGTGGAAGGTAAACAGAATCTGGTGATTATGGGTAGGAAAACCTGGTTC
 2761 TCCATTCTTGAGAAGAATCGACCTTTAAAGGACAGAATTAATATAGTTCTCAGTAGAGAA
 2821 CTCAAAGAACCACACGAGGAGCTCATTTTCTTGCCAAAAGTTTGGATGATGCCTTAAGA
 2881 CTTATTGAACAACCGGAATTGGCAAGTAAAGTAGACATGGTTTGGATAGTCGGAGGCAGT
 2941 TCTGTTTACCAGGAAGCCATGAATCAACCAGGCCACCTCAGACTCTTTGTGACAAGGATC
 3001 ATGCAGGAATTTGAAAGTGACACGTTTTTCCCAGAAATTGATTTGGGGAAATATAAACTT
 3061 CTCCCAGAATACCCAGGCGTCTCTCTGAGGTCCAGGAGGAAAAAGGCATCAAGTATAAG
 3121 TTTGAAGTCTACGAGAAGAAAGACTAACAGGAAGATGCTTTCAGTTCTCTGCTCCCCTC
 3181 CTAAAGCTATGCATTTTATAAGACCATGGGACTTTTGCTGGCTTT**AGATCT**TTGTGAAG
 3241 GAACCTTACTTCTGTGGTGTGACATAATTGGACAACTACCTACAGAGATTTAAAGCTCT
 3301 AAGGTAAATATAAAATTTTTAAGTGTATAATGTGTTAACTACTGATTCTAATTGTTTGT
 3361 GTATTTTAGATTCCAACCTATGGAAGTATGAATGGGAGCAGTGGTGGAATGCCTTTAAT
 3421 GAGGAAAACCTGTTTTGCTCAGAAGAAATGCCATCTAGTGATGATGAGGCTACTGCTGAC
 3481 TCTCAACATTCTACTCCTCCAAAAAAGAAGAGAAAGGTAGAAGACCCCAAGGACTTTCCT
 3541 TCAGAATTGCTAAGTTTTTTGAGTCATGCTGTGTTTAGTAATAGAACTCTTGCTTGCTTT
 3601 GCTATTTACACCACAAAGGAAAAAGCTGCACTGCTATACAAGAAAATTATGGAAAAATAT
 3661 TCTGTAACCTTTATAAGTAGGCATAACAGTTATAATCATAACATACTGTTTTTCTTACT
 3721 CCACACAGGCATAGAGTGTCTGCTATTAATAACTATGCTCAAAAATTGTGTACCTTTAGC
 3781 TTTTAAATTTGTAAAGGGTTAATAAGGAATATTTGATGTATAGTGCCTTGACTAGAGAT
 3841 CATAATCAGCCATACCACATTTGTAGAGGTTTTACTTGCTTTAAAAAACCTCCCACACCT
 3901 CCCCCTGAACCTGAAACATAAAATGAATG**CAATTG**TTGTTGTTAACTTGTTTATTGCAGC
 3961 TTATAATGGTTACAAATAAAGCAATAGCATCACAAATTCACAAATAAAGCATTTTTTTC
 4021 ACTGCATTCTAGTTGTGGTTTGTCCAACTCATCAATGTATCTTATCATGTCTGGATCTA
 4081 ATAAAAGATATTTATTTTCATTAGATATGTGTGTTGGTTTTTTGTGTGCAGTGCCTCTAT
 4141 CTGGAGGCCAGGTAGGGCTGGCCTTGGGGGAGGGGAGGCCAGAATGACTCCAAGAGCTA
 4201 CAGGAAGGCAGGTCAGAGACCCCACTGGACAAACAGTGGCTGGACTCTGCACCATAACAC

Fig. 27 14

4261 ACAATCAACAGGGGAGTGAGCTGGAAATTTGCTAGC**GAATTC**cagcacactggcgggccgt
Spe I
4321 **tACTAGT**TATTAATAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATATGGAGTT
4381 CCGCGTTACATA**AACTT**ACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCCGCCC
4441 ATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACG
4501 TCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATAT
4561 GCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCA
SnaB I
4621 GTACATGACCTTATGGGACTTTCTACTTGGCAGTACATC**TACGTAT**TAGTCATCGCTAT
4681 TACCATGGTGATGCGGTTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTTGACTCACG
4741 GGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCA
4801 ACGGGACTTTCCAAAATGTCGTAACAACTCCGCCCCATTGACGCAAATGGGCGGTAGGCG
4861 TGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTCAGATCGCCTGGAG
4921 ACGCCATCCACGCTGTTTTGACCTCCATAGAAGACACCGGGACCGATCCAGCCTCCGCGG
4981 CCGGGAACGGTGCATTGGAACGCGGATTCCCCGTGCCAAGAGTGACGTAAGTACCGCCTA
5041 TAGAGTCTATAGGCCACCCCCCTTGGCTTCTTATGCATGCTATACTGTTTTTGGCTTGGG
Bpu1102I
5101 GTCTATACACCCCCGCTTCCTCATGTTATAGGTGATGGTATAG**CTTAGC**CTATAGGTGTG
Xcm I
5161 GGTTATTGACCATTATTG**CCACTCCCCCTATTGGT**GACGATACTTTCCATTACTAATCCA
5221 TAACATGGCTCTTTGCCACAACCTCTCTTTATTGGCTATATGCCAATACACTGTCCTTCAG
5281 AGACTGACACGGACTCTGTATTTTTACAGGATGGGGTCTCATTATTATTACAAATTCA
5341 CATATACAACACCACCGTCCCCAGTGCCCGCAGTTTTTTATTAAACATAACGTGGGATCTC
BspE I
5401 CACGCGAATCTCGGGTACGTGT**TCCGGA**CATGGGCTCTTCTCCGGTAGCGGCGGAGCTTC
5461 TACATCCGAGCCCTGCTCCCATGCCTCCAGCGACTCATGGTCGCTCGGCAGCTCCTTGCT
5521 CCTAACAGTGGAGGCCAGACTTAGGCACAGCACGATGCCACCACCACCAGTGTGCCGCA
.....
5581 CAAGGCCGTGGCGGTAGGGTATGTGTCTGAAAATGAGCTCggggagcgggcttgaccgc
(Pvu II)
5641 tgacgcatttgaagacttaaggcagcggcagaagaagatgcagg**cagctg**agttgttgt
5701 gttctgataagagtcagaggtaactccggttgcggtgctgttaacgggtggagggcagtg
5761 agtctgagcagtactcgttgctgccgcgcgcgccaccagacataatagctgacagactaa
Mlu I
5821 cagactgttcctttccatgggtccttttctgcagtcaccgtccttgac**ACGCGT**CTCGGGA
Hind III
5881 **AGCTT**GCCGCCACCATGGGATGGAGCTGGGTCTTTCTCTTTCTCCTGTCAGGAAGTGCAG
M G W S W V F L F L L S G T A

Fig. 27 /5

(Pvu II)

5941 GTGTCCTCTCTGAGGTC**CAGCTG**CAACAGTCTGGACCTGAGCTGGTGAAGCCTGGGGCTT
 G V L S E V Q L Q Q S G P E L V K P G A

Xba I Dra III

6001 CAGTAAAGATGTCCTGCAAGACT**TCTAGATA**CACATTCACTGAATACACCATA**CACTGGG**
 S V .K M S C K T S R Y T F T E Y T I H W

CDR 1

6061 **TG**AGACAGAGCCATGGAAAGAGCCTTGAGTGGATTGGAGGTATTAATCCTAACAATGGTA
 V R Q S H G K S L E W I G G I N P N N G

6121 TTCCTAACTACAACCAGAAGTTCAAGGGCAGGGCCACATTGACTGTAGGCAAGTCCTCCA
 I P N Y N Q K F K G R A T L T V G K S S

CDR 2

6181 GCACCGCCTACATGGAGCTCCGCAGCCTGACATCTGAGGATTCTGCGGTCTATTTCTGTG
 S T A Y M E L R S L T S E D S A V Y F C

6241 CAAGAAGAAGAATCGCCTATGGTTACGACGAGGGCCATGCTATGGACTACTGGGGTCAAG
 A R R R I A Y G Y D E G H A M D Y W G Q

CDR 3 BamH I

6301 GAACCTCAGTCACCGTCTCCTCAGGTGAGT**GGATCC**TCTGCGCCTGGGCCCAGCTCTGTC
 G T S V T V S S

6361 CCACACCGCGGTACATGGCACCACCTCTCTTGCAGCCTCCACCAAGGGCCCATCGGTCT
 S T K G P S V

6421 TCCCCCTGGCACCCCTCCTCCAAGAGCACCTCTGGGGGCACAGCGGCCCTGGGCTGCCTGG
 F P L A P S S K S T S G G T A A L G C L

Age I

6481 TCAAGGACTACTTCCCCGA**ACCGGT**GACGGTGTCTGTGGAAGTCAAGGCGCCCTGACCAGCG
 V K D Y F P E P V T V S W N S G A L T S

6541 GCGTGCACACCTTCCCGGCTGTCCTACAGTCCTCAGGACTCTACTCCCTCAGCAGCGT**GG**
 G V H T F P A V L Q S S G L Y S L S S V

BstE II

6601 **TGACCGT**GCCCTCCAGCAGCTTGGGCACCCAGACCTACATCTGCAACGTGAATCACAAGC
 V T V P S S S L G T Q T Y I C N V N H K

6661 CCAGCAACACCAAGGTGGACAAGAAAGTTGAGCCCCAAATCTTGTGACAAAACCTCACACAT
 P S N T K V D K K V E P K S C D K T H T

6721 GCCCACCGTGCCCAGCACCTGAACTCCTGGGGGGACCGTCAGTCTTCTCTTCCCCCAA
 C P P C P A P E L L G G P S V F L F P P

6781 AACCCAAGGACACCCTCATGATCTCCCGGACCCCTGAGGTCACATGCGTGGTGGTGGACG
 K P K D T L M I S R T P E V T C V V V D

6841 TGAGCCACGAAGACCCTGAGGTCAAGTTCAACTGGTACGTGGACGGCGTGGAGGTGCATA
 V S H E D P E V K F N W Y V D G V E V H

6901 ATGCCAAGACAAAGCCGCGGGAGGAGCAGTACAACAGCACGTACCGGGTGGTCAGCGTCC
 N A K T K P R E E Q Y N S T Y R V V S V

6961 TCACCGTCCTGCACCAGGACTGGCTGAATGGCAAGGAGTACAAGTGCAAGGTCTCCAACA
 L T V L H Q D W L N G K E Y K C K V S N

Fig. 27 /6

7021 AAGCCCTCCCAGCCCCCATCGAGAAAACCATCTCCAAAGCCAAAGGGCAGCCCCGAGAAC
 K A L P A P I E K T I S K A K G Q P R E
 7081 CACAGGTGTACACCCTGCCCCCATCCCGGGAGGAGATGACCAAGAACCAGGTCAGCCTGA
 P Q V Y T L P P S R E E M T K N Q V S L
 7141 CCTGCCTGGTCAAAGGCTTCTATCCCAGCGACATCGCCGTGGAGTGGGAGAGCAATGGGC
 T C L V K G F Y P S D I A V E W E S N G
 7201 AGCCGGAGAACAACACTACAAGACCACGCCTCCCGTGCTGGACTCCGACGGCTCCTTCTTCC
 Q P E N N Y K T T P P V L D S D G S F F
 7261 TCTACAGCAAGCTCACCGTGGACAAGAGCAGGTGGCAGCAGGGGAACGTCTTCTCATGCT
 L Y S K L T V D K S R W Q Q G N V F S C
 7321 CCGTGATGCATGAGGCTCTGCACAACCACTACACGCAGAAGAGCCTCTCCCTGTCTCCGG
 S V M H E A L H N H Y T Q K S L S L S P
 Ngom I
 7381 GTAAATGAGTGCACGGCCGGCAAGCCCCGCTCCCCGGGCTCTCGCGGTGCACGAGGAT
 G K *
 7441 GCTTGGCACGTACCCCCTGTACATACTTCCCGGGCGCCAGCATGGAAATAAAGCACCGG
 7501 ATCTAATAAAAGATATTTATTTTCATTAGATATGTGTGTTGGTTTTTTTGTGTGCAGTGCC
 7561 TCTATCTGGAGGCCAGGTAGGGCTGGCCTTGGGGGAGGGGGAGGCCAGAATGACTCCAAG
 7621 AGCTACAGGAAGGCAGGTCAGAGACCCCACTGGACAAACAGTGGCTGGACTCTGCACCAT
 7681 AACACACAATCAACAGGGGAGTGAGCTGGaaatttgctagcgaattaattc 7731

Fig. 28:

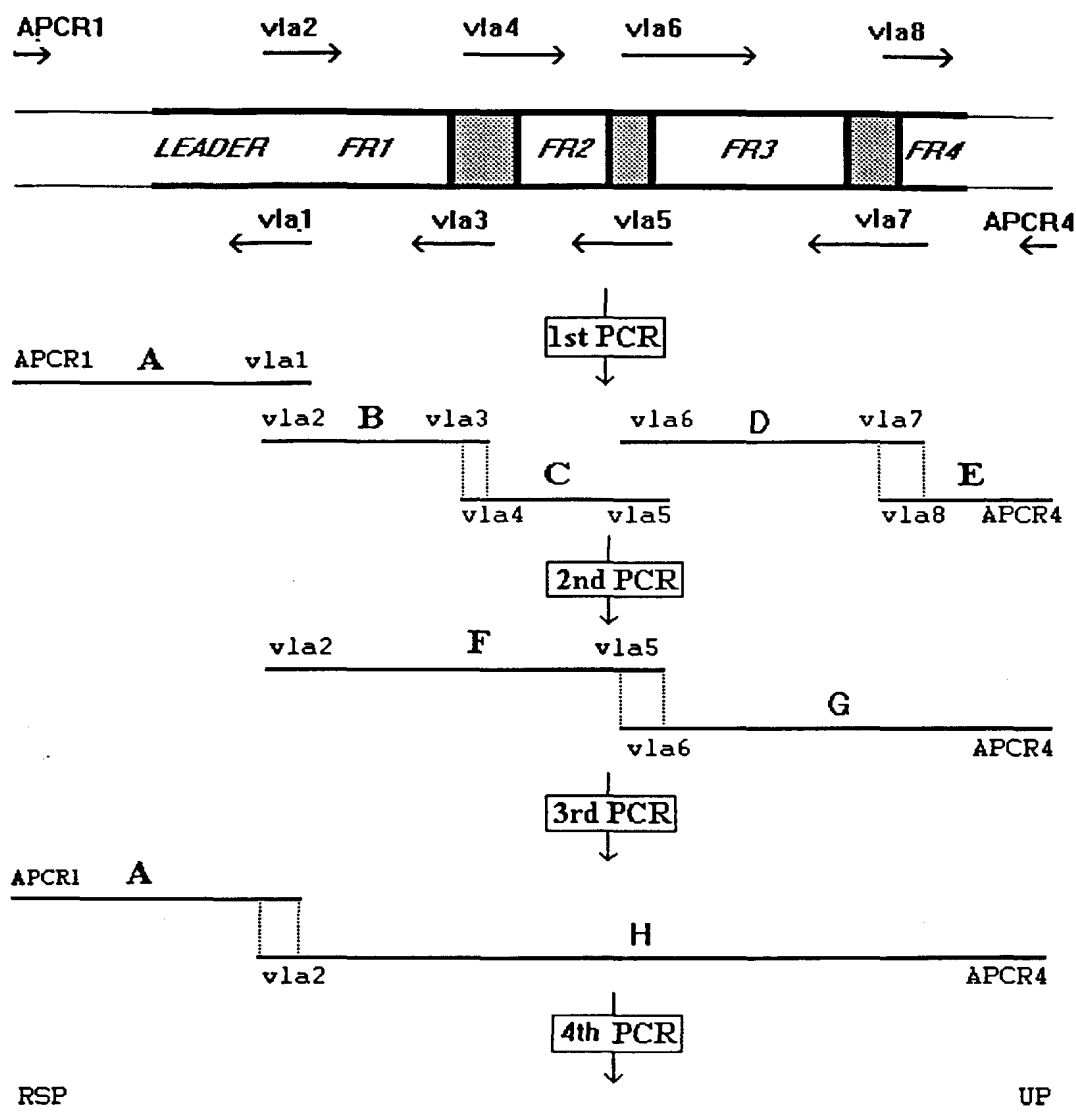


Fig. 29 /1

	1	D	I	V	M	T	Q	S	P	D	S	L	A	V	S	L	G	E	R	19	A				
A		GAC	ATT	GTG	ATG	ACC	CAA	TCT	CCA	GAC	TCT	TTG	GCT	GTG	TCT	CTA	GGG	GAG	AGG	GCC					
B						
C		---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---					
	20	T	I	N	C	CDR1										27	A	B	C	D	E	F	28		32
A		ACC	ATC	AAC	TGC	AAG	TCC	AGT	CAG	AGC	CTT	TTA	TAT	TCT	AGA	AAT	CAA	AAG	AAC	TAC					
B						
C		---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---					
	33	L	A	W	Y	Q	Q	K	P	G	Q	P	P	K	L	L	I	F	W	A	51				
A		TTG	GCC	TGG	TAT	CAG	CAG	AAA	CCA	GGA	CAG	CCA	CCC	AAA	CTC	CTC	ATC	TTT	TGG	GCT					
B		.	.	.	F					
C		---	---	---	-TC	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---					
	52	S	T	R	E	S	G	V	P	D	R	F	S	G	S	G	F	G	T	D	70				
A		AGC	ACT	AGG	GAA	TCT	GGG	GTA	CCT	GAT	AGG	TTC	AGT	GGC	AGT	GGG	TTT	GGG	ACA	GAC					
B						
C		---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---					
	71	F	T	L	T	I	S	S	L	Q	A	E	D	V	A	V	Y	Y	C	88					
A		TTC	ACC	CTC	ACC	ATT	AGC	AGC	CTG	CAG	GCT	GAA	GAT	GTG	GCA	GTT	TAT	TAC	TGT						
B		D	.						
C		---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	G	---						
	89	Q	Q	Y	F	S	Y	P	L	T	F	G	Q	G	T	K	V	E	I	K	107				
A		CAG	CAA	TAT	TTT	AGC	TAT	CCG	CTC	ACG	TTC	GGA	CAA	GGG	ACC	AAG	GTG	GAA	ATA	AAA					
B		---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---					
C		---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---					

Fig. 30 /1

1 gaattccagc acactggcgg ccgttACTAG TTATTAATAG TAATCAATTA
 Spe I
 51 CGGGGTCATT AGTTCATAGC CCATATATGG AGTTCGCGT TACATAACTT
 101 ACGGTAAATG GCCCGCCTGG CTGACCGCCC AACGACCCCC GCCCATTGAC
 151 GTCAATAATG ACGTATGTTC CCATAGTAAC GCCAATAGGG ACTTTCCATT
 201 GACGTCAATG GGTGGAGTAT TTACGGTAAA CTGCCCACTT GGCAGTACAT
 251 CAAGTGTATC ATATGCCAAG TACGCCCCCT ATTGACGTCA ATGACGGTAA
 301 ATGGCCCGCC TGGCATTATG CCCAGTACAT GACCTTATGG GACTTTCCTA
 351 CTTGGCAGTA CATCTACGTA TTAGTCATCG CTATTACCAT GGTGATGCGG
 SnaB I
 401 TTTTGGCAGT ACATCAATGG GCGTGGATAG CGGTTTGA CT CACGGGGATT
 451 TCCAAGTCTC CACCCCAT TG ACGTCAATGG GAGTTTGTTT TGGCACCAAA
 501 ATCAACGGGA CTTTCCAAAA TGTCGTAACA ACTCCGCCCC ATTGACGCAA
 551 ATGGGCGGTA GCGTGTACG GTGGGAGGTC TATATAAGCA GAGCTCGTTT
 601 AGTGAACCGT CAGATCGCCT GGAGACGCCA TCCACGCTGT TTTGACCTCC
 651 ATAGAAGACA CCGGGACCGA TCCAGCCTCC GCGGCCGGGA ACGGTGCATT
 Sac II
 701 GGAACGCGGA TTCCCCGTGC CAAGAGTGAC GTAAGTACCG CCTATAGAGT
 751 CTATAGGCCC ACCCCCTTGG CTTCTTATGC ATGCTATACT GTTTTTGGCT
 801 TGGGGTCTAT ACACCCCCGC TTCCTCATGT TATAGGTGAT GGTATAGCTT
 851 AGCCTATAGG TGTGGGTTAT TGACCATTAT TGACCACTCC CCTATTGGTG
 901 ACGATACTTT CCATTACTAA TCCATAACAT GGCTCTTTGC CACAACTCTC
 951 TTTATTGGCT ATATGCCAAT AACTGTCCT TCAGAGACTG ACACGGACTC
 1001 TGTATTTT TA CAGGATGGGG TCTCATTTAT TATTTACAAA TTCACATATA
 1051 CAACACCACC GTCCCCAGTG CCCGCAGTTT TTATTAAACA TAACGTGGGA
 (BspE I)
 1101 TCTCCACGCG AATCTCGGGT ACGTGTTCCG GACATGGGCT CTTCTCCGGT
 1151 AGCGGCGGAG CTTCTACATC CGAGCCCTGC TCCATGCCT CCAGCGACTC
 1201 ATGGTCGCTC GGCAGCTCCT TGCTCCTAAC AGTGGAGGCC AGACTTAGGC

Fig. 30 /2

1251 ACAGCACGAT GCCCACCACC ACCAGTGTGC CGCACAAGGC CGTGGCGGTA

1301 GGGTATGTGT CTGAAAATGA GCTCqggggaq cgggcttgca ccgctgacgc

Afl II

1351 atttgggaaga ctttaaggcag cggcagaaga agatgcaggc agctgagttg

1401 ttgtgttctg ataagagtcg gaggtaacctc ccgttgcggt gctgttaacg

1451 gtggagggca gtgtagtctg agcagtactc gttgctgccg cgcgcgccac

1501 cagacataat agctgacaga ctaacagact gttcctttcc atggggtcttt

Mlu I

Hind III

1551 tctgcagtcac cgcctctga cagcgcctc gggaaagcttG CCGCCACCAT

M

1601 GGAGACAGAC ACACTCCTGC TATGGGTGCT GCTGCTCTGG GTTCCAGGTT

E T D T L L L W V L L L W V P G

(BspE I)

1651 CCTCCGGA GA CATTGTGATG ACCCAATCTC CAGACTCTTT GGCTGTGTCT

S S G D I V M T Q S P D S L A V S

1701 CTAGGGGAGA GGGCCACCAT CAACTGCAAG TCCAGTCAGA GCCTTTTATA

L G E R A T I N C K S S Q S L L Y

Xbal

CDR 1

1751 TTCTAGAAAT CAAAAGAACT ACTTGGCCTG GTATCAGCAG AAACCAGGAC

SRN QKN YLAW YQQ KPG

1801 AGCCACCCAA ACTCCTCATC TTTTGGGCTA GCACTAGGGA ATCTGGGGTA **KpnI**

Q P P K L L I F W A S T R E S G V

CDR 2

1851 CCTGATAGGT TCAGTGGCAG TGGGTTTGGG ACAGACTTCA CCCTCACCAT

P D R E S G S G F G T D F T L T I

1901 TAGCAGCCTG CAGGCTGAAG ATGTGGCAGT TTATTACTGT CAGCAATATT

SSLQAE DVAVYYCQQY

1951 TTAGCTATCC GCTCACGTTC GGACAAGGGA CCAAGGTGGA AATAAACGT

F S Y P L T F G Q G T K V E I K R

CDR 3

BamH I

2001 GAGT**ggatcc** ATCTGGGATA AGCATGCTGT TTTCTGTCTG TCCCTAACAT

2051 GCCCTGTGAT TATGCGCAAA CAACACACCC AAGGGCAGAA CTTTGTTACT

2101 TAAACACCAT CCTGTTTGCT TCTTTCCTCA GGAAGTGTGG CTGCACCATC

T V A A P S

2151 TGTCTTCATC TTCCCGCCAT CTGATGAGCA GTTGAAATCT GGAAGTGCCT

V F I F P P S D E Q L K S G T A

2201 CTGTTGTGTG CCTGCTGAAT AACTTCTATC CCAGAGAGGC CAAAGTACAG

S V V C L L N N F Y P R E A K V Q

2251 TGG AAGGTGG ATAACGCCCT CCAATCGGGT AACTCCCAGG AGAGTGTCAC

W K V D N A L Q S G N S Q E S V T

2301 AGAGCAGGAC AGCAAGGACA GCACCTACAG CCTCAGCAGC ACCCTGACGC

Fig. 30 /3

E Q D S K D S T Y S L S S T L T
 2351 TGAGCAAAGC AGACTACGAG AAACACAAAG TCTACGCCTG CGAAGTCACC
 L S K A D Y E K H K V Y A C E V T
 2401 CATCAGGGCC TGAGCTCGCC CGTCACAAAG AGCTTCAACA GGGGAGAGTG
 H Q G L S S P V T K S F N R G E C
 2451 TTAGAGGGAG AAGTGCCCCC ACCTGCTCCT CAGTTCCAGC CTGACCCCCCT
 * Psp5 II
 2501 CCCATCCTTT GGCCTCTGAC CCTTTTTCCA CAGGGGACCT ACCCCTATTG
 2551 CGGTCCTCCA GCTCATCTTT CACCTCACCC CCCTCCTCCT CCTTGGCTTT
 2601 AATTATGCTA ATGTTGGAGG AGAATGAATA AATAAAGTGA ATCTTTGCAC
 2651 CTGTGGTGGA TCTAATAAAA GATATTTATT TTCATTAGAT ATGTGTGTTG
 2701 GTTTTTTGTG TGCAGTGCCT CTATCTGGAG GCCAGGTAGG GCTGGCCTTG
 2751 GGGGAGGGGG AGGCCAGAAT GACTCCAAGA GCTACAGGAA GGCAGGTCAG
 2801 AGACCCCACT GGACAAACAG TGGCTGGACT CTGCACCATA ACACACAATC
 2851 AACAGGGGAG TGAGCTGGAA ATTTGCTAGC GAATTCTTGA AGACGAAAGG
 2901 GCCTCGTGAT ACGCCTATTT TTATAGGTTA ATGTCATGAT AATAATGGTT
 2951 TCTTAGACGT CAGGTGGCAC TTTTCGGGGA AATGTGCGCG GAACCCCTAT
 3001 TTGTTTATTT TTCTAAATAC ATTCAAATAT GTATCCGCTC ATGAGACAAT
 3051 AACCCTGATA AATGCTTCAA TAATATTGAA AAAGGAAGAG TATGAGTATT
 3101 CAACATTTCC GTGTCGCCCT TATTCCCTTT TTTGCGGCAT TTTGCCTTCC
 3151 TGTTTTTGCT CACCCAGAAA CGCTGGTGAA AGTAAAAGAT GCTGAAGATC
 3201 AGTTGGGTGC ACGAGTGGGT TACATCGAAC TGGATCTCAA CAGCGGTAAG
 3251 ATCCTTGAGA GTTTTCGCCC CGAAGAACGT TTTCCAATGA TGAGCACTTT
 3301 TAAAGTTCTG CTATGTGGCG CGGTATTATC CCGTGTTGAC GCCGGGCAAG
 3351 AGCAACTCGG TCGCCGCATA CACTATTCTC AGAATGACTT GGTGAGTAC
 3401 TCACCAGTCA CAGAAAAGCA TCTTACGGAT GGCATGACAG TAAGAGAATT
 3451 ATGCAGTGCT GCCATAACCA TGAGTGATAA CACTGCGGCC AACTTACTTC
 Pvu I
 3501 TGACAACGAT CGGAGGACCG AAGGAGCTAA CCGCTTTTTT GCACAACATG
 3551 GGGGATCATG TAACTCGCCT TGATCGTTGG GAACCGGAGC TGAATGAAGC

Fig. 30 /4

3601 CATACCAAAC GACGAGCGTG ACACCACGAT GCCTGCAGCA ATGGCAACAA
 3651 CGTTGCGCAA ACTATTA ACT GGC GAACTAC TTA CTCTAGC TTCCCGGCAA
 3701 CAATTAATAG ACTGGATGGA GCGGATAAA GTTGCAGGAC CACTTCTGCG
 3751 CTCGGCCCTT CCGGCTGGCT GGTTTATTGC TGATAAATCT GGAGCCGGTG
 3801 AGCGTGGGTC TCGCGGTATC ATTGCAGCAC TGGGGCCAGA TGGTAAGCCC
 3851 TCCCGTATCG TAGTTATCTA CACGACGGGG AGTCAGGCAA CTATGGATGA
 3901 ACGAAATAGA CAGATCGCTG AGATAGGTGC CTC ACTGATT AAGCATTGGT
 3951 AACTGTCAGA CCAAGTTTAC TCATATATAC TTTAGATTGA TTAAAACTT
 4001 CATTTTTAAT TAAAAGGAT CTAGGTGAAG ATCCTTTTTG ATAATCTCAT
 4051 GACCAAATC CCTTAACGTG AGTTTTCGTT CCACTGAGCG TCAGACCCCG
 4101 TAGAAAAGAT CAAAGGATCT TCTTGAGATC CTTTTTTTCT GCGCGTAATC
 4151 TGCTGCTTGC AAACAAAAA ACCACCGCTA CCAGCGGTGG TTTGTTTGCC
 4201 GGATCAAGAG CTACCAACTC TTTTCCGAA GGTA ACTGGC TTCAGCAGAG
 4251 CGCAGATACC AAATACTGTC CTTCTAGTGT AGCCGTAGTT AGGCCACCAC
 4301 TTCAAGAACT CTGTAGCACC GCCTACATAC CTCGCTCTGC TAATCCTGTT
 4351 ACCAGTGGCT GCTGCCAGTG GCGATAAGTC GTGTCTTACC GGGTTGGACT
 4401 CAAGACGATA GTTACCGGAT AAGGCGCAGC GGTCGGGCTG AACGGGGGGT
 4451 TCGTGACAC AGCCAGCTT GGAGCGAACG ACCTACACCG AACTGAGATA
 4501 CCTACAGCGT GAGCTATGAG AAAGCGCCAC GCTTCCCGAA GGGAGAAAGG
 4551 CGGACAGGTA TCCGGTAAGC GGCAGGGTCG GAACAGGAGA GCGCACGAGG
 4601 GAGCTTCCAG GGGGAAACGC CTGGTATCTT TATAGTCCTG TCGGGTTTCG
 4651 CCACCTCTGA CTTGAGCGTC GATTTTTGTG ATGCTCGTCA GGGGGGCGGA
 4701 GCCTATGGAA AAACGCCAGC AACGCGGCCT TTTACGGTT CCTGGCCTTT
 4751 TGCTGGCCTT TTGCTCACAT GTTCTTTCCT GCGTTATCCC CTGATTCTGT
 4801 GGATAACCGT ATTACCGCCT TTGAGTGAGC TGATACCGCT CGCCGCAGCC

BspLU11I

Fig. 30 /5

4851 GAACGACCGA GCGCAGCGAG TCAGTGAGCG AGGAAGCGGA AGAGCGCCTG
 4901 ATGCGGTATT TTCTCCTTAC GCATCTGTGC GGTATTTTAC ACCGCATATG
 4951 GTGCACTCTC AGTACAATCT GCTCTGATGC CGCATAGTTA AGCCAGTATA
 5001 CACTCCGCTA TCGCTACGTG ACTGGGTCAT GGCTGCGCCC CGACACCCGC
 5051 CAACACCCGC TGACGCGCCC TGACGGGCTT GTCTGCTCCC GGCATCCGCT
 5101 TACAGACAAG CTGTGACCGT CTCCGGGAGC TGCATGTGTC AGAGGTTTTTC
 5151 ACCGTCATCA CCGAAACGCG CGAGGCAGCT GTGGAATGTG TGTCAGTTAG
 5201 GGTGTGGAAG GTCCCCAGGC TCCCCAGCAG GCAGAAGTAT GCAAAGCATG
 5251 CATCTCAATT AGTCAGCAAC CAGGCTCCCC AGCAGGCAGA AGTATGCAA
 5301 GCATGCATCT CAATTAGTCA GCAACCATAG TCCCGCCCCT AACTCCGCCC
 5351 ATCCCGCCCC TAACTCCGCC CAGTTCCGCC CATTCTCCGC CCCATGGCTG
 5401 ACTAATTTTT TTTATTTATG CAGAGGCCGA GGCCGCCTCG GCCTCTGAGC
 5451 TATTCCAGAA GTAGTGAGGA GGCTTTTTTG GAGGCCTAGG CTTTTGCAA
 5501 AAGCTAGCTT CACGCTGCCG CAAGCACTCA GGGCGCAAGG GCTGCTAAAG
 5551 GAAGCGGAAC ACGTAGAAAG CCAGTCCGCA GAAACGGTGC TGACCCCGGA
 5601 TGAATGTCAG CTAAGGGCT ATCTGGACAA GGGAAAACGC AAGCGCAAAG
 5651 AGAAAGCAGG TAGCTTGCAAG TGGGCTTACA TGGCGATAGC TAGACTGGGC
 5701 GGTTTTATGG ACAGCAAGCG AACCAGGAATT GCCAGCTGGG GCGCCCTCTG
 5751 GTAAGGTTGG GAAGCCCTGC AAAGTAACT GGATGGCTTT CTTGCCGCCA
 5801 AGGATCTGAT GCGCAGGGG ATCAAGATCT GATCAAGAGA CAGGATGAGG
 5851 ATCGTTTCGC ATGATTGAAC AAGATGGATT GCACGCAGGT TCTCCGGCCG
 5901 CTTGGGTGGA GAGGCTATTC GGCTATGACT GGGCACAACA GACAATCGGC
 5951 TGCTCTGATG CCGCCGTGTT CCGGCTGTCA GCGCAGGGGC GCCCGGTTCT
 6001 TTTTGTCAAG ACCGACCTGT CCGGTGCCCT GAATGAACTG CAGGACGAGG
 6051 CAGCGCGGCT ATCGTGGCTTG GCCACGACGG GCGTTCCTTG CGCAGCTGTG

Fig. 30 /6

6101 CTCGACGTTG TCACTGAAGC GGGAAGGGAC TGGCTGCTAT TGGGCGAAGT
 6151 GCCGGGGCAG GATCTCCTGT CATCTCACCT TGCTCCTGCC GAGAAAGTAT
 6201 CCATCATGGC TGATGCAATG CGGCGGCTGC ATACGCTTGA TCCGGCTACC
 6251 TGCCCATTCG ACCACCAAGC GAAACATCGC ATCGAGCGAG CACGTACTCG
 6301 GATGGAAGCC GGTCTTGTCG ATCAGGATGA TCTGGACGAA GAGCATCAGG
 6351 GGCTCGCGCC AGCCGAAGT TCGCCAGGC TCAAGGCGCG CATGCCCGAC
 6401 GGCGAGGATC TCGTCGTGAC CCATGGCGAT GCCTGCTTGC CGAATATCAT
 6451 GGTGGAAAAT GGCCGCTTTT CTGGATTCAT CGACTGTGGC CGGCTGGGTG
 6501 TGGCGGACCG CTATCAGGAC ATAGCGTTGG CTACCCGTGA TATTGCTGAA
 6551 GAGCTTGCGG GCGAATGGGC TGACCGCTTC CTCGTGCTTT ACGGTATCGC
 6601 CGCTCCCGAT TCGCAGCGCA TCGCCTTCTA TCGCCTTCTT GACGAGTTCT
 6651 TCTGAGCGGG ACTCTGGGGT TCGAAATGAC CGACCAAGCG ACGCCCAACC
 6701 TGCCATCACG AGATTTCGAT TCCACCGCCG CCTTCTATGA AAGGTTGGGC
 6751 TTCGGAATCG TTTTCCGGGA CGCCGGCTGG ATGATCCTCC AGCGCGGGGA
 6801 TCTCATGCTG GAGTTCTTCG CCCACCCCGG GCTCGATCCC CTCGCGAGTT
 6851 GGTTCACTG CTGCCTGAGG CTGGACGACC TCGCGGAGTT CTACCGGCAG
 6901 TGCAAATCCG TCGGCATCCA GAAACCAGC AGCGGCTATC CGCGCATCCA
 6951 TGCCCCCGAA CTGCAGGAGT GGGGAGGCAC GATGGCCGCT TTGGTCCCGG
 7001 ATCTTTGTGA AGGAACCTTA CTTCTGTGGT GTGACATAAT TGGACAAACT
 7051 ACCTACAGAG ATTTAAAGCT CTAAGGTAAA TATAAAATTT TTAAGTGTAT
 7101 AATGTGTAA ACTACTGATT CTAATTGTTT GTGTATTTTA GATTCCAACC
 7151 TATGGAAGT ATGAATGGGA GCAGTGGTGG AATGCCTTTA ATGAGGAAAA
 7201 CCTGTTTTGC TCAGAAGAAA TGCCATCTAG TGATGATGAG GCTACTGCTG
 7251 ACTCTCAACA TTCTACTCCT CAAAAAAGA AGAGAAAGGT AGAAGACCCC
 7301 AAGGACTTTC CTTCAGAATT GCTAAGTTTT TTGAGTCATG CTGTGTTTAG

Fig. 30 /7

7351 TAATAGAACT CTTGCTTGCT TTGCTATTTA CACCACAAAG GAAAAAGCTG
 7401 CACTGCTATA CAAGAAAATT ATGGAAAAAT ATTCTGTAAC CTTTATAAGT
 7451 AGGCATAACA GTTATAATCA TAACATACTG TTTTTTCTTA CTCCACACAG
 7501 GCATAGAGTG TCTGCTATTA ATAACTATGC TCAAAAATTG TGTACCTTTA
 7551 GCTTTTTAAT TTGTAAAGGG GTTAATAAGG AATATTTGAT GTATAGTGCC
 7601 TTGACTAGAG ATCATAATCA GCCATACCAC ATTTGTAGAG GTTTTACTTG
 7651 CTTTAAAAAA CCTCCACAC CTCCCCCTGA ACCTGAAACA TAAATGAAT
 Mun I
 7701 GCAATTGTTG TTGTAACTT GTTTATTGCA GCTTATAATG GTTACAAATA
 7751 AAGCAATAGC ATCACAAATT TCACAAATAA AGCATTTTTT TCACTGCATT
 7801 CTAGTTGTGG TTTGTCCAAA CTCATCAATG TATCTTATCA TGTCTGGATC
 7851 TAATAAAAGA TATTTATTTT CATTAGATAT GTGTGTTGGT TTTTGTGTG
 7901 CAGTGCCTCT ATCTGGAGGC CAGGTAGGGC TGGCCTTGGG GGAGGGGGAG
 7951 GCCAGAATGA CTCCAAGAGC TACAGGAAGG CAGGTCAGAG ACCCCACTGG
 8001 ACAAACAGTG GCTGGACTCT GCACCATAAC ACACAATCAA CAGGGGAGTG
 8051 AGCTGGAAAT TTGCTAGC

Fig. 31

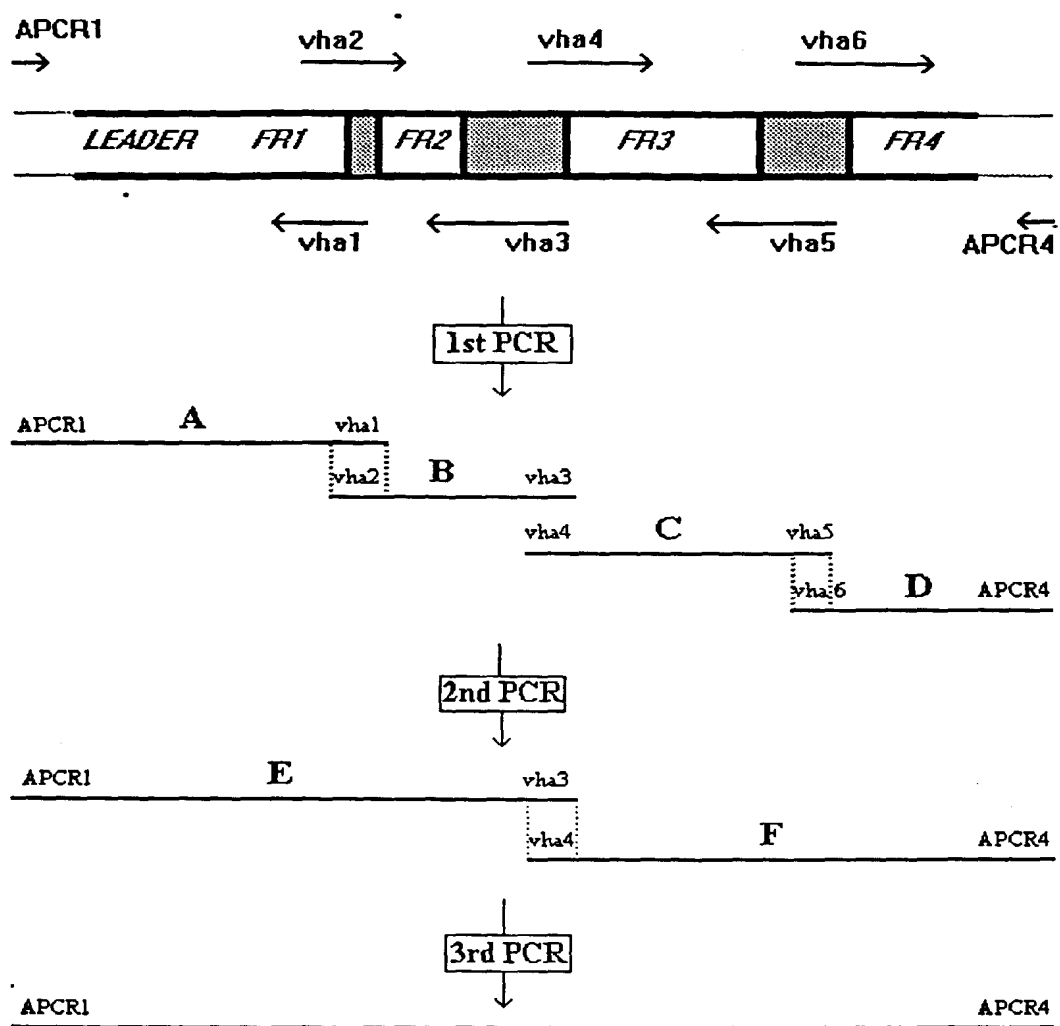


Fig. 32 /1

	1									10									19	
A	Q	V	Q	L	V	Q	S	G	A	E	V	K	K	P	G	A	S	V	K	
	CAG	GTG	CAA	CTA	GTG	CAG	TCC	GGC	GCC	GAA	GTG	AAG	AAA	CCC	GGT	GCT	TCC	GTG	AAA	
B	
C	
D	
E	
	20										30	CDR1							38	
A	V	S	C	K	T	S	R	Y	T	F	T	E	Y	T	I	H	W	V	R	
	GTC	AGC	TGT	AAA	ACT	AGT	AGA	TAC	ACC	TTC	ACT	GAA	TAC	ACC	ATA	CAC	TGG	GTT	AGA	
B	
C	
D	
E	G	
							G													
	39										49	52					A	53		56
A	Q	A	P	G	Q	R	L	E	W	I	G	G	I	N	P	N	N	G	I	
	CAG	GCC	CCT	GGC	CAA	AGG	CTG	GAG	TGG	ATA	GGA	GGT	ATT	AAT	CCT	AAC	AAT	GGT	ATT	
B	
C	
D	
E	
	57	CDR2										70							75	
A	P	N	Y	N	Q	K	F	K	G	R	A	T	L	T	V	G	K	S	A	
	CCT	AAC	TAC	AAC	CAG	AAG	TTC	AAG	GGC	CGG	GCC	ACC	TTG	ACC	GTA	GGC	AAG	TCT	GCC	
B	
C	V	.	I	.	.	D	T	.	.	
D	V	.	I	.	.	D	T	.	.	
E	V	.	I	.	.	D	T	.	.	
											T		A-C			A	CC			
											T		A-C			A	CC			
											T		A-C			A	CC			
	76							82	A	B	C	83							91	
A	S	T	A	Y	M	E	L	S	S	L	R	S	E	D	T	A	V	Y	Y	
	AGC	ACC	GCC	TAC	ATG	GAA	CTG	TCC	AGC	CTG	CGC	TCC	GAG	GAC	ACT	GCA	GTC	TAC	TAC	
B	T	
C	
D	F.	
E	T	

Fig. 32 /2

	92				CDR3					100	A	B	C	D	I	J	K	101		103
	C	A	R		R	R	I	A	Y	G	Y	D	E	G	H	A	M	D	Y	W
A	TGC	GCC	AGA		AGA	AGA	ATC	GCC	TAT	GGT	TAC	GAC	GAG	GGC	CAT	GCT	ATG	GAC	TAC	TGG
B
C	---	---	---		---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
D
E	---	---	---		---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---

	104										113
	G	Q	G	T	L	V	T	V	S	S	
A	GGT	CAA	GGA	ACC	CTT	GTC	ACC	GTC	TCC	TCA	
B	
C	---	---	---	---	---	---	---	---	---	---	
D	
E	---	---	---	---	---	---	---	---	---	---	

Fig. 33 /1

1 TTGAAGACGAAAGGGCCTCGTGATACGCCTATTTTTATAGGTTAATGTCATGATAATAAT
 61 GGTTTCTTAGACGTCAGGTGGCACTTTTCGGGGAAATGTGCGCGGAACCCCTATTTGTTT
 121 ATTTTTCTAAATACATTCAAATATGTATCCGCTCATGAGACAATAACCCTGATAAATGCT
 181 TCAATAATATTGAAAAAGGAAGAGTATGAGTATTCAACATTTCCGTGTCGCCCTTATTCC
 241 CTTTTTTGCGGCATTTTGCCTTCCTGTTTTTGTCTACCCAGAAACGCTGGTGAAAGTAAA
 301 AGATGCTGAAGATCAGTTGGGTGCACGAGTGGGTTACATCGAACTGGATCTCAACAGCGG
 361 TAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTCCAATGATGAGCACTTTTAAAGT
 421 TCTGCTATGTGGCGCGGTATTATCCCGTGTTGACGCCGGGCAAGAGCAACTCGGTGCGCG
 481 CATACACTATTCTCAGAATGACTTGTTGAGTACTCACCAGTCACAGAAAAGCATCTTAC
 541 GGATGGCATGACAGTAAGAGAATTATGCAGTGCTGCCATAACCATGAGTGATAACACTGC
 601 GGCCAACTTACTTCTGACAACGATCGGGAGGACCGAAGGAGCTAACCGCTTTTTTGCACAA
 661 CATGGGGGATCATGTAACTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACC
 721 AAACGACGAGCGTGACACCAGATGCCTGCAGCAATGGCAACAACGTTGCGCAAACTATT
 781 AACTGGCGAACTACTTACTCTAGCTTCCCGCAACAATTAATAGACTGGATGGAGGCGGA

Fig. 33 /2

841 TAAAGTTGCAGGACCACTTCTGCGCTCGGCCCTTCCGGCTGGCTGGTTTATTGCTGATAA
 901 ATCTGGAGCCGGTGAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAA
 961 GCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGGATGAACGAAA
 1021 TAGACAGATCGCTGAGATAGGTGCCTCACTGATTAAGCATTGGTAACTGTCAGACCAAGT
 1081 TTACTCATATATACTTTAGATTGATTTAAACTTCATTTTTTAATTTAAAAGGATCTAGGT
 1141 GAAGATCCTTTTTTGATAATCTCATGACCAAAATCCCTTAACGTGAGTTTTTCGTTCCACTG
 1201 AGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTTCTGCGCGT
 1261 AATCTGCTGCTTGCAAACAAAAAAACCACCGCTACCAGCGGTGGTTTGTGTGCGGATCA
 1321 AGAGCTACCAACTCTTTTTCCGAAGGTAAGTGGCTTCAGCAGAGCGCAGATACCAAATAC
 1381 TGTCTTCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTAC
 1441 ATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCT
 1501 TACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTGGGGCTGAACGGG
 1561 GGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCTACA
 1621 GCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGT
 1681 AAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGGTA
 1741 TCTTTATAGTCCTGTGCGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTC
 1801 GTCAGGGGGGCGGAGCCTATGGAAAACGCCAGCAACGCGGCCTTTTTACGGTTCCTGGC
 1861 CTTTTGCTGGCCTTTTGCTCACATGTTCCTTTCCTGCGTTATCCCTGATTCTGTGGATAA
 1921 CCGTATTACCGCCTTTGAGTGAGCTGATACCGCTCGCCGCAGCCGAACGACCGAGCGCAG
 1981 CGAGTCAGTGAGCGAGGAAGCGGAAGAGCGCCTGATGCGGTATTTCTCCTTACGCATCT
 2041 GTGCGGTATTTACACCGCATATGGTGCACTCTCAGTACAATCTGCTCTGATGCCGCATA
 2101 GTTAAGCCAGGTATACACTCCGCTATCGCTACGTGACTGGGTCATGGCTGCGCCCCGACAC
 2161 CCGCCAACACCCGCTGACGCGCCCTGACGGGCTTGTCTGCTCCCGGCATCCGCTTACAGA
 2221 CAAGCTGTGACCGTCTCCGGGAGCTGCATGTGTGAGAGGTTTTACCGTTCATCACCAGAA
 2281 CGCGCGAGGCAGCATGCATCTCAATTAGTCAGCAACCATAGTCCCGCCCCTAACCTCCGCC
 2341 CATCCCGCCCCTAACCTCCGCCAGTTCCGCCCATTTCTCCGCCCATGGCTGACTAATTTT
 2401 TTTTATTTATGCAGAGGCCGAGGGCCGCCTCGGCCTCTGAGCTATTCCAGAAGTAGTGAGG
 2461 AGGCTTTTTTGAGGCCTAGGCTTTTGCAAAAAGCTAGCTTACAGCTCAGGGCTGCGATT

Fig. 33 /2

2521 TCGCGCCAAACTTGACGGCAATCCTAGCGTGAAGGCTGGTAGGATTTTATCCCCGCTGCC
 2581 ATCATGGTTCGACCATTTGAACTGCATCGTCGCCGTGTCCCAAATATGGGGATTGGCAAG
 2641 AACGGAGACCTACCCTGGCCTCCGCTCAGGAACGAGTTCAAGTACTTCCAAAGAATGACC
 2701 ACAACCTCTTCAGTGAAGGTAAACAGAATCTGGTGATTATGGGTAGGAAAACCTGGTTC
 2761 TCCATTCTCTGAGAAGAATCGACCTTTAAAGGACAGAATTAATATAGTTCTCAGTAGAGAA
 2821 CTCAAAGAACCACCACGAGGAGCTCATTTTCTTGCCAAAAGTTTGGATGATGCCTTAAGA
 2881 CTTATTGAACAACCGGAATTGGCAAGTAAAGTAGACATGGTTTGGATAGTCGGAGGCAGT
 2941 TCTGTTTACCAGGAAGCCATGAATCAACCAGGCCACCTCAGACTCTTTGTGACAAGGATC
 3001 ATGCAGGAATTTGAAAGTGACACGTTTTTCCCAGAAATTGATTTGGGGAAATATAAACTT
 3061 CTCCCAGAATACCCAGGCGTCTCTCTGAGGTCCAGGAGGAAAAAGGCATCAAGTATAAG
 3121 TTTGAAGTCTACGAGAAGAAAGACTAACAGGAAGATGCTTTCAAGTTCTCTGCTCCCCCTC
 3181 CTAAAGCTATGCATTTTTTATAAGACCATGGGACTTTTGCTGGCTTT**AGATCT**TTGTGAAG
 3241 GAACCTTACTTCTGTGGTGTGACATAATTGGACAAACTACCTACAGAGATTTAAAGCTCT
 3301 AAGGTAAATATAAAATTTTTAAGTGTATAATGTGTAAACTACTGATTCTAATTGTTTGT
 3361 GTATTTTAGATTCCAACCTATGGAACCTGATGAATGGGAGCAGTGGTGGAAATGCCTTTAAT
 3421 GAGGAAAACCTGTTTTGCTCAGAAGAAATGCCATCTAGTGATGATGAGGCTACTGCTGAC
 3481 TCTCAACATTCTACTCCTCCAAAAAAGAAGAGAAAGGTAGAAGACCCCAAGGACTTTCCT
 3541 TCAGAAATGCTAAGTTTTTTGAGTCATGCTGTGTTTAGTAATAGAACTCTTGCTTGCTTT
 3601 GCTATTTACACCACAAAGGAAAAAGCTGCACTGCTATACAAGAAAATTATGGAAAAATAT
 3661 TCTGTAACCTTTATAAGTAGGCATAACAGTTATAATCATAACATACTGTTTTTTCTTACT
 3721 CCACACAGGCATAGAGTGTCTGCTATTAATAACTATGCTCAAAAATTGTGTACCTTTAGC
 3781 TTTTAAATTTGTAAAGGGGTTAATAAGGAATATTTGATGTATAGTGCCTTGACTAGAG**AT**
 3841 **CATAATC**AGCCATACCACATTTGTAGAGGTTTTACTTGCTTTAAAAAACCTCCCACACCT
 3901 CCCCCTGAACCTGAAACATAAAATGAATG**CAATTG**TTGTTGTTAACTGTTTATTGCAGC
 3961 TTATAATGGTTACAAATAAAGCAATAGCATCACAAATTTACAAATAAAGCATTTTTTTC
 4021 ACTGCATTCTAGTTGTGGTTTGTCCAAACTCATCAATGTATCTTATCATGTCTGGATCTA
 4081 ATAAAAGATATTTATTTTCATTAGATATGTGTGTTGGTTTTTTGTGTGCAGTGCCTCTAT
 4141 CTGGAGGCCAGGTAGGGCTGGCCTTGGGGGAGGGGAGGCCAGAATGACTCCAAGAGCTA

Bgl II

BsaB I

Mun I

Fig. 33 /3

4201 CAGGAAGGCAGGTCAGAGACCCCACTGGACAAACAGTGGCTGGACTCTGCACCATAACAC
EcoR I
4261 ACAATCAACAGGGGAGTGAGCTGGAAATTTGCTAGC**GAATTC**cagcacactggcgggccgt
(Spe I)
4321 t**ACTAGT**TATTAATAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATATGGAGTT
4381 CCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCCGCCC
4441 ATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACG
4501 TCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATAT
4561 GCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCA
SnaB I
4621 GTACATGACCTTATGGGACTTTCTACTTGGCAGTACATC**TACGTA**TTAGTCATCGCTAT
4681 TACCATGGTGATGCGGTTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTTGACTCACG
4741 GGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAATCA
4801 ACGGGACTTTCCAAAATGTCGTAACAACCTCCGCCCCATTGACGCAAATGGGCGGTAGGCG
4861 TGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTCAGATCGCCTGGAG
4921 ACGCCATCCACGCTGTTTTGACCTCCATAGAAGACACCGGGACCGATCCAGCCTCCGCGG
4981 CCGGGAACGGTGCAATTGGAACGCGGATTCCCCGTGCCAAGAGTGACGTAAGTACCGCCTA
5041 TAGAGTCTATAGGCCCACCCCTTGGCTTCTTATGCATGCTATACTGTTTTTGGCTTGGG
Bpu1102I
5101 GTCTATACACCCCGCTTCCTCATGTTATAGGTGATGGTATAG**GCTTAGC**CTATAGGTGTG
Xcm I
5161 GGTTATTGACCATTATTGAC**CCACTCCCCTATTGGT**GACGATACTTTCCATTACTAATCCA
5221 TAACATGGCTCTTTGCCACAACCTCTCTTTATTGGCTATATGCCAATACACTGTCCTTCAG
5281 AGACTGACACGGACTCTGTATTTTTTACAGGATGGGGTCTCATTTATTATTACAAATTCA
5341 CATATACAACACCACCGTCCCCAGTGCCCGCAGTTTTTTATTAAACATAACGTGGGATCTC
BspE I
5401 CACGCGAATCTCGGGTACGTGT**TCCGGAC**ATGGGCTCTTCTCCGGTAGCGGCGGAGCTTC
5461 TACATCCGAGCCCTGCTCCCATGCCCTCCAGCGACTCATGGTCGCTCGGCAGCTCCTTGCT
5521 CCTAACAGTGAGAGGCCAGACTTAGGCACAGCACGATGCCACCACCACAGTGTGCCGCA
5581 CAAGGCCGTGGCGGTAGGGTATGTGTCTGAAAATGAGCTCggggagcgggcttgaccgc
(Pvu II)
5641 tgacgcatttgaagacttaaggcagcggcagaagaagatgcagg**cagctg**agttgttgt
5701 gttctgataagagtcagaggttaactccggttgcggtgctgttaacggtggagggcagtgt
5761 agtctgagcagtagctcggtgctgcccgcgcgcgccaccagacataatagctgacagactaa
Mlu I
5821 cagactgttcctttccatgggtctttttctgcagtcaccgtccttgac**ACGCGTCTCGGGA**

Fig. 33 /5

6961 TCACCGTCCTGCACCAGGACTGGCTGAATGGCAAGGAGTACAAGTGCAAGGTCTCCAACA
 L T V L H Q D W L N G K E Y K C K V S N
 7021 AAGCCCTCCCAGCCCCCATCGAGAAAACCATCTCCAAAGCCAAAGGGCAGCCCCGAGAAC
 K A L P A P I E K T I S K A K G Q P R E
 7081 CACAGGTGTACACCCTGCCCCCATCCCGGGAGGAGATGACCAAGAACCAGGTCTCAGCCTGA
 P Q V Y T L P P S R E E M T K N Q V S L
 7141 CCTGCCTGGTCAAAGGCTTCTATCCCAGCGACATCGCCGTGGAGTGGGAGAGCAATGGGC
 T C L V K G F Y P S D I A V E W E S N G
 7201 AGCCGGAGAACTACAAGACCACGCCTCCCGTGCTGGACTCCGACGGCTCCTTCTTCC
 Q P E N N Y K T T P P V L D S D G S F F
 7261 TCTACAGCAAGCTCACCGTGGACAAGAGCAGGTGGCAGCAGGGGAACGTCTTCTCATGCT
 L Y S K L T V D K S R W Q Q G N V F S C
 7321 CCGTGATGCATGAGGCTCTGCACAACCACTACACGCAGAAGAGCCTCTCCCTGTCTCCGG
 S V M H E A L H N H Y T Q K S L S L S P
 Ngm I
 7381 GTAAATGAGTGCACGGCCGGCAAGCCCCGCTCCCCGGGCTCTCGCGGTGCACGAGGAT
 G K *
 7441 GCTTGGCACGTACCCCCTGTACATACTTCCCGGGCGCCAGCATGGAAATAAAGCACCGG
 7501 ATCTAATAAAAGATATTTATTTTCATTAGATATGTGTGTTGGTTTTTTGTGTGCAGTGCC
 7561 TCTATCTGGAGGCCAGGTAGGGCTGGCCTTGGGGGAGGGGGAGGCCAGAATGACTCCAAG
 7621 AGCTACAGGAAGGCAGGTCAGAGACCCCACTGGACAAACAGTGGCTGGACTCTGCACCAT
 7681 AACACACAATCAACAGGGGAGTGAGCTGGaaatttgctagcgaattaattc 7731

Fig. 34 A

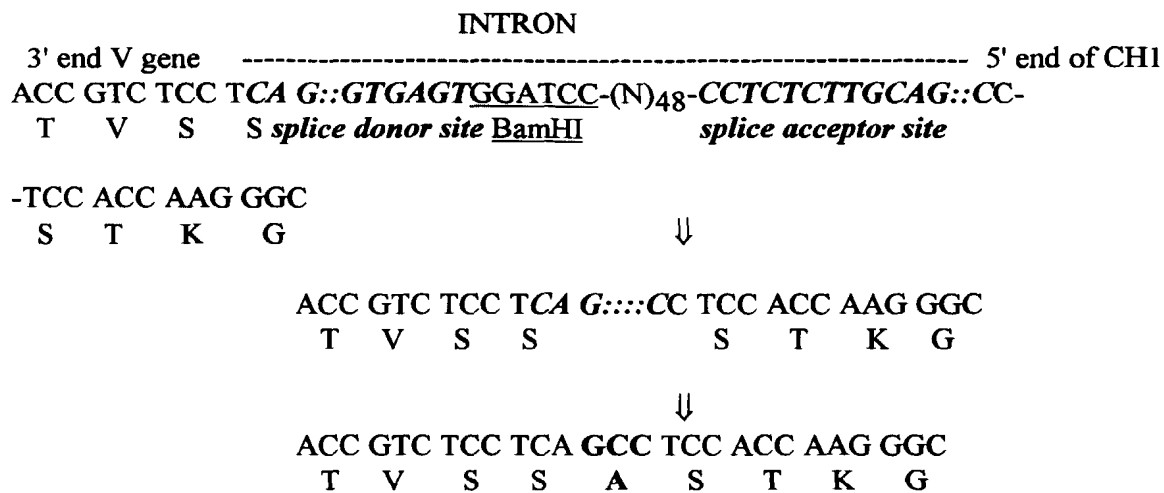
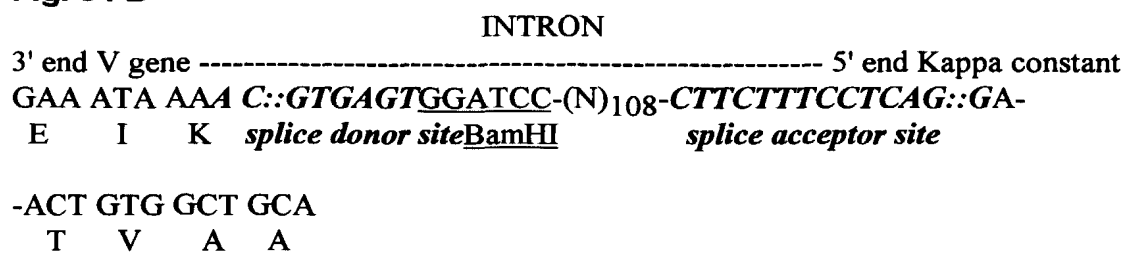
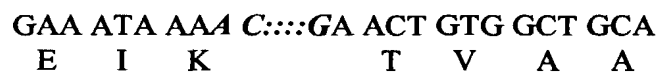


Fig. 34 B

⇓



⇓

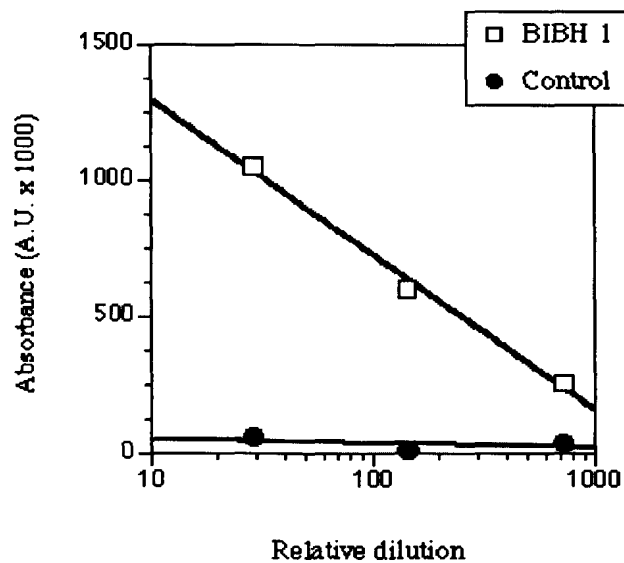
**Fig. 35**

Fig. 36

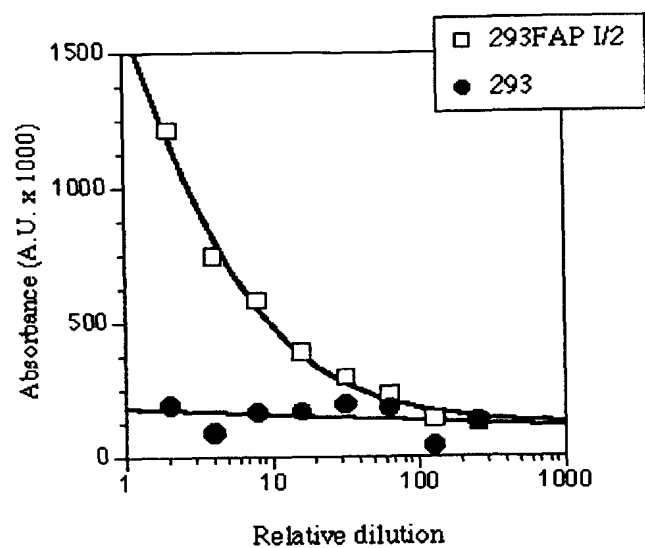
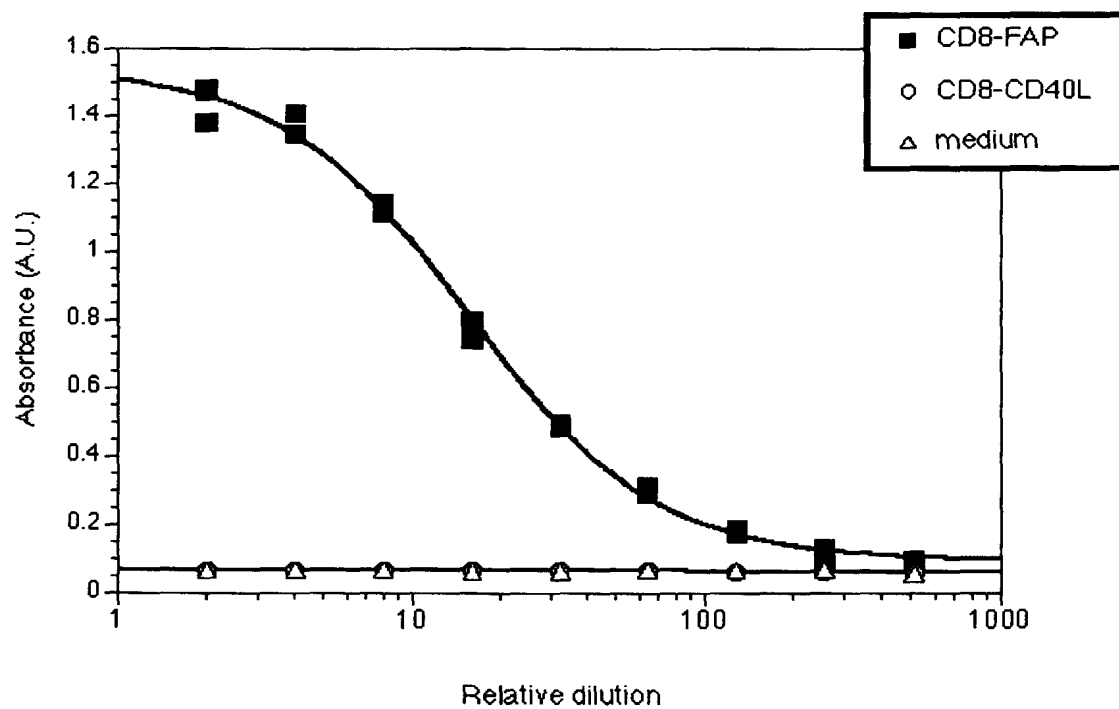


Fig. 37





European Patent
Office

PARTIAL EUROPEAN SEARCH REPORT

Application Number

which under Rule 45 of the European Patent Convention shall be considered, for the purposes of subsequent proceedings, as the European search report

DOCUMENTS CONSIDERED TO BE RELEVANT			
Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (Int.Cl.6)
Y	WELT ET AL.: "Antibody targeting in metastatic colon cancer: a phase I study of monoclonal antibody F19 against a cell-surface protein of reactive tumor stromal fibroblasts" JOURNAL OF CLINICAL ONCOLOGY, vol. 12, no. 6, June 1994, pages 1193-1203, XP002088696 * abstract * * page 1193, column 1, line 1 - page 1194, column 2, line 4 * * page 1202, column 2, paragraph 2 * ---	1-65	C12N15/13 C07K16/40 C07K16/46 C12N15/62 C12N15/85 C12N5/10 C07K19/00 A61K47/48 A61K51/10 A61K39/395 G01N33/577 G01N33/574
Y	WO 93 05804 A (SLOAN KETTERING INST CANCER) 1 April 1993 * abstract; claims 1-23 * ---	1-65	
Y	US 5 693 761 A (SCHNEIDER WILLIAM P ET AL) 2 December 1997 * abstract * * examples 3-9 * * column 2, line 36 - column 3, line 59 * --- -/--	1-65	
			TECHNICAL FIELDS SEARCHED (Int.Cl.6)
			C07K
INCOMPLETE SEARCH			
<p>The Search Division considers that the present application, or one or more of its claims, does/do not comply with the EPC to such an extent that a meaningful search into the state of the art cannot be carried out, or can only be carried out partially, for these claims.</p> <p>Claims searched completely :</p> <p>Claims searched incompletely :</p> <p>Claims not searched :</p> <p>Reason for the limitation of the search:</p> <p>see sheet C</p>			
Place of search		Date of completion of the search	Examiner
MUNICH		21 December 1998	Muller-Thomalla, K
CATEGORY OF CITED DOCUMENTS		<p>T : theory or principle underlying the invention</p> <p>E : earlier patent document, but published on, or after the filing date</p> <p>D : document cited in the application</p> <p>L : document cited for other reasons</p> <p>& : member of the same patent family, corresponding document</p>	
<p>X : particularly relevant if taken alone</p> <p>Y : particularly relevant if combined with another document of the same category</p> <p>A : technological background</p> <p>O : non-written disclosure</p> <p>P : intermediate document</p>			

EPO FORM 1503 03/82 (P04-C07)



European Patent
Office

PARTIAL EUROPEAN SEARCH REPORT

Application Number
EP 98 10 7925

DOCUMENTS CONSIDERED TO BE RELEVANT			CLASSIFICATION OF THE APPLICATION (Int.Cl.6)
Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	
Y	STUDNICKA G M ET AL: "Human-engineered monoclonal antibodies retain full specific binding activity by preserving non- CDR complementarity-modulating residues." PROTEIN ENGINEERING, (1994 JUN) 7 (6) 805-14. JOURNAL CODE: PR1. ISSN: 0269-2139., XP000447301 ENGLAND: United Kingdom * page 805, column 1, line 1 - page 806, column 2, paragraph 1 * * page 808, column 2, paragraph 2 - page 812, column 1, paragraph 1 * * page 813, column 2, paragraph 1 *	1-65	
Y	WRIGHT A ET AL: "Genetically engineered antibodies: progress and prospects." CRITICAL REVIEWS IN IMMUNOLOGY, (1992) 12 (3-4) 125-68. REF: 252 JOURNAL CODE: AF1. ISSN: 1040-8401., XP000616488 United States * page 139, column 2, paragraph 3 - page 141, column 1, paragraph 3 * * page 157, column 2, paragraph 3 - page 158, column 1, paragraph 1 *	1-65	TECHNICAL FIELDS SEARCHED (Int.Cl.6)
A	WO 94 05690 A (SMITHKLINE BEECHAM CORP ;US ARMY (US); GROSS MITCHELL STUART (US);) 17 March 1994 * claim 5; figure 3 *	14-17	

EPO FORM 1503 03.82 (P04C10)



European Patent
Office

**INCOMPLETE SEARCH
SHEET C**

Application Number
EP 98 10 7925

Although claims 50-52,54,55,57,61,62,65 are directed to a method of treatment of the human/animal body and/or a diagnostic method practised on the human/animal body (Article 52(4) EPC), the search has been carried out and based on the alleged effects of the compound/composition.